Essays on the Psychology of Economic Choice

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktorin der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

Fanny Brun
from Lancy, Genève

Approved in February 2019 at the request of
Prof. Dr. Roberto A. Weber
Prof. Dr. Michel Maréchal

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 13.02.2019

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgements

Writing my doctoral thesis has been a beautiful challenge, and I would not have achieved it without the support of several persons.

First and foremost, I would like to wholeheartedly thank my supervisor Roberto Weber for his remarkable support in the last three years. I belong to the lucky PhD students who had the chance to learn so much from their mentor. Thank you, Roberto, for showing me how to run projects methodically, with precision, rigor (and a touch of humor), and for constantly pushing me towards the excellence. You helped me to find my way in life by discovering the topic that deeply moves me. I am very grateful for that. I would also like to thank you for being not only a professor but a human being who always takes time to enquire about the well-being of your team members, and for the overall pleasant working atmosphere that you generated. I always enjoyed a lot the nice lunch breaks all together, with lots of laughs and of positive energy. I will miss this pleasant environment.

I would also like to thank my second supervisor, Michel Maréchal, for having been highly accessible and ready to help when I needed. Your advices about my last research project and your support in the preparation of jobs interviews were very insightful.

A big thank you goes also to all brilliant and wonderful members of Roberto's team. In particular, I would like to thank you, Florian, for being such a great co-author, for your readiness to help, for your good mood and for your love for cats (two cats for you: =^-^= and ^-.-^). Thank you, Vanessa, for your so positive attitude, your continuous listening and for generating a beautiful harmony in the group. Thank you, Lea, for your great friendship and your precious advices in all domains of life. Thank you, Yi-Shan, for your kindness and your willingness to help. And finally, another big thank you to Marina, for your fantastic energy, your optimism, your unfailing support and for generating such a great working atmosphere.

I would also like to express my deep gratitude to a couple of colleagues and friends without whom my PhD experience would have looked totally different and from whom I learned a lot on a professional as well as on a personal level.

Dear Juliette and Max, I am so lucky that we started together this adventure. You were always there: supporting me in hard times, and celebrating with me in happy times. You both have always impressed me by your intellectual and social intelligence, and your inspiring way of living your life. Your constant and contagious optimism and positive energy empowered me a lot during these four years. A deep thank you for all that.

Dear Caro, we started to really know each other in the second year of PhD. And I am so glad we did. I discovered an extremely bright woman, with an outstanding emotional intelligence and capacity to listen. Thank you for your unfailing intellectual and moral support and for all the funny moments that we shared. I am very pleased and lucky to have met you, and I can't wait to continue this beautiful friendship!

A special thought goes also to Christian. Thank you for your listening, for the many energizing lunch breaks and for the fascinating discussions. I would also like to mention a couple of additional persons who have been also very important to me over the course of my PhD studies: a big thank you to May-ma-Jacotte, Andi, Vera, Gabrielle, Brilé, Anna, Tania as well as to my grand-mother who always believed in me.

Last but not least, I am very thankful to the UBS center that supported me financially in the last four years of my doctoral studies and for all the inspiring events and meetings that I had the chance to attend.

Contents

Tables

Figures

Chapter 1 – Introduction

In my dissertation, I look at several psychological factors affecting the utility of decision makers. Having a strong interest in social and environmental causes, I particularly focus on ethical and socially responsible decisions of consumers (chapters 2 to 4) and of workers (chapter 5).

The first two chapters of my thesis investigate the issue of food waste. Since most of this issue is generated at the household level, the purpose of this research is to better understand what leads individuals to discard edible food. Chapter 2 looks at the psychological mechanism behind people's food waste practices, while chapter 3 studies what types of personality traits can best predict such behaviors.

Chapter 2 presents the results of a lab-in-the-field experiment jointly developed with Prof. Roberto Weber where we investigate the psychological effect of freshness information on people's experienced utility of consumption. Previous research finds that date labels play an important role in individuals' decision to throw away food. People's gustatory experience of food was also shown to decrease along the shelf-life of food and to drop at the expiry date. However, these existing studies have several weaknesses and none investigates the *psychological mechanism* explaining people's food waste practices. In particular, it remains unclear whether individuals dispose of old food only for safety reasons or also for freshness ones. In our study, we disentangle those two possible drivers of food waste behavior. We investigate the effects of information on the (lack of) freshness of a food item on individuals' gustatory experience, provided that the food is safe to eat – a phenomenon that we call *psychological depreciation*.

Chapter 3 is based on the survey data collected in the experiment presented in chapter 2. In this work, I look at how certain individuals' psychological factors relate to their food waste practices. More precisely, I look at whether risk preferences (in general, in the food

and in the health/safety domains), descriptive norms (how much food individuals believe people usually throw away), and personal attitudes towards food waste (more specifically, how important it is for individuals to minimize the amount of food that they discard) can predict individuals' reported tendency to dispose of edible food. In a second step, I investigate which of these predictors performs best. Finally, I also look at the behavioral consistency of participants self-declared food waste practices by testing whether their *reported* behavior significantly correlates with their *actual* one in the laboratory.

Overall, the findings of these two studies provide a better understanding of individuals' behaviors towards "old" food and thus useful insights to policymakers aiming to effectively address the issue of food waste. Our results may also be of interest for actors in the food industry who would like for instance to reduce the amount of unsold perishable food items.

Chapter 4 looks at consumers' happiness with their choices. In the last century, the offer of products and services, as well as the number of possible life options (e.g. where to live, what job to choose, what education to follow) have dramatically increased, leading individuals to encounter a growing number of decisions every day. In this context, it becomes particularly relevant to better understand what drives people's satisfaction (or lack thereof) with their choices. In this paper, Reto Odermatt, Itay Sisso, and I look at how the choice set impacts individuals' satisfaction with their choice. We hypothesize that even after having chosen one's preferred option, the opportunity costs formed by the rejected alternative (in a choice set consisting of two goods) may diminish the post-choice utility one derives from the selected good. We also suggest that the size of opportunity costs matters: rejecting an attractive alternative may be more detrimental for post-choice utility than rejecting a less desirable one. We denote the extent to which opportunity costs matter for satisfaction by *psychological opportunity costs* (POCs); *Psychological* since we expect these opportunity costs to have a negative impact on people's *post-decision* satisfaction.

This concept contradicts the standard economic view that opportunity costs are only relevant for the decision-making process and do not matter anymore once the decision has been made. Our results allow to better understand what can enhance or diminish consumers' happiness with their choices, and may therefore be of interest not only for companies aiming to retain their customers, but also for governments that want to increase the well-being of their population.

In the last chapter of my dissertation, Florian Schneider, Prof. Roberto Weber and I study immoral labor markets and workers' psychological aversion to perform immoral jobs. We define a job as *immoral* if it can potentially cause harm to society. In the last decades, several corporate scandals made public the existence of such behaviors within firms. For instance, the 2008 financial crisis revealed that bankers had intentionally sold toxic assets to unsuspecting clients (US Department of Justice, 2016). Also, tobacco companies were accused of using marketing strategies to mislead the public about the harmful effects of smoking (Heath, 2016) or to encourage youth smoking (Bates and Rowell, 1998). Besides, there are also jobs that *inherently* involve immoral acts by their nature (e.g. manufacturing weapons, marketing cigarettes, or engaging in predatory lending). While there is a widespread perception that immoral jobs pay a wage premium to compensate for workers' aversion to act immorally, there is only suggestive evidence for this mechanism. Also, no previous studies looked at whether heterogeneity in willingness to perform unethical work leads to differential labor market outcomes. In Chapter 5, we attempt to address these questions using both laboratory experiments and survey data.

Chapter 2 – Psychological Depreciation and Food Waste

Joint with Roberto A. Weber[1]

2.1 Introduction

Imagine that every household, once per day, prepared a nutritious meal and immediately threw it in the trash without taking a bite. While this is perhaps a shocking example, the reality is that households regularly do precisely this without noticing.

Around one third of the food produced worldwide is lost or wasted every year. This represents about 1.3 billion tons of discarded edible food (FAO, 2011). For example, in Switzerland the equivalent of 320 grams of food, i.e. almost a whole meal, is discarded every day for every inhabitant (WWF, 2012). Food waste is also harmful to the environment—food production is responsible for one third of all greenhouse gas emissions (Garnett, 2011), and is highly resource-intensive (e.g. water, energy, land etc., FAO, 2013; Godfray et al., 2010; NRDC, 2012). Food waste is also costly from households' perspective: in developed countries, households spend 7% of their income on food, 30% of which is wasted (FOA, 2016). This represents an annual loss estimated between $1,350 and $2,275 for a typical American four-person household (NRDC, 2012). The loss of edible food takes place throughout the whole food supply chain (FAO, 2011), but an important part of it (from 52% for Europe to 61% for North America and Oceania, in percent of kcal wasted) takes place at the household level (Lipinski et al., 2013).

While there are many causes of food waste, explicit freshness informational cues—such as date labels—seem to play an important role. In 2012, the Swiss federal office for the environment analyzed the content of citizens' trash cans to discover that 20% of total waste

[1] Please cite as Brun, Fanny and Roberto A. Weber (2018) "Psychological Depreciation and Food Waste," Working Paper.

was unopened packaged food that had passed the expiry date (SFOE, 2012). In the United Kingdom, about 27% of household waste is due to food spoilage from not being used in time (WRAP, 2015). While it is certainly reasonable to discard hazardous or inedible food, paying too much attention to the date labels when deciding whether to eat or discard a product may be suboptimal. Indeed, date labels are loosely regulated and may provide limited information about actual food quality. For instance, in the USA or in Switzerland, sellers of food items are free to set the expiry dates as they want. However, they must be fit for consumption (for the USA: NRDC, 2013; for Switzerland: DFI, 2017), otherwise they can be held responsible if it caused health issues. As a consequence, expiry dates are usually conservative and quite noisy signals of quality and safety. Moreover, expiry dates confound two possible kinds of inferences that consumers may draw about a food product, since a proximate or past expiry date can indicate both that a product is unsafe or that it has lower levels of quality or freshness.

We report an experiment that attempts to hold as much possible constant about the actual quality and safety of a food item, to isolate whether informational cues about freshness alone lead to the decision to forgo eating food and thus contribute to food waste. In particular, we are interested in investigating whether knowledge that a food item is "new" versus "old" influences consumers' perceptions of the food's quality and their experiences when eating the food. We focus our attention on people's perception of food freshness with respect to how long an item has been stored. This is important, as the time that has passed since a food product was made or purchased is often an easily accessible cue to people evaluating whether to eat or discard the food. Moreover, an expired food item is not only unfresh but also potentially harmful for one's health. By focusing on old but unexpired products, we rule out the potential negative effect of people's safety concerns on their

gustatory experience. Our findings thus address whether freshness concerns alone lead people to discard food.

In our study, we measure people's experienced utility of consuming a food item in a controlled and real-choice situation, where we vary by treatment the provision of information on the production date. More precisely, we test if the knowledge of a production date *far in the past* affects individuals' gustatory experience *negatively*, a phenomenon that we call *psychological depreciation.* Psychological depreciation is potentially important for understanding food waste, as it may lead people to make suboptimal decisions and produce avoidable waste. That is, if people find food less appealing simply because it is older, they may dispose of it more readily based on cues about how fresh it is, even when its objective qualities would still make it desirable—i.e., when the food's *actual* depreciation is much smaller than its *psychological* depreciation. Finding evidence for the existence—or the absence—of this phenomenon would provide useful insights on how to set up effective interventions to combat food waste.

In our lab-in-the-field study, we distribute halves of packaged sandwiches to students on a university campus over the lunch break. Participants state how eager they are to eat their sandwich before starting it. Then while eating, they fill out a survey where they indicate, among others, the taste of the sandwich and their willingness-to-pay (WTP hereafter) for receiving the second half of the exact same sandwich. They have to return the completed survey and any remaining unfinished portions of the sandwich to receive their participation fee. If they are willing to pay at a randomly drawn price, this price is deducted from the participation fee and they receive the second half of their sandwich.

This experiment is a 2 x 2 between-subjects design, where we vary whether the production date is close (*Fresh*) or far (*Old*) in the past, and whether the production date is disclosed (*Salient*) or hidden (*Blind*). We chose a specific type of packaged sandwich with

a long storage duration for which quality is likely to remain stable over time. Importantly, participants were always informed that their sandwich is safe to eat (the expiry date was always respected). Although we intentionally selected a good for which we expected the quality would remain fairly stable over time, we allow the possibility that participants in the *Blind* conditions will enjoy the old sandwich to a slightly lesser extent than a fresh one, reflecting the possibility of actual depreciation of sandwich quality over several additional days of storage. However, the main hypothesis we test—of psychological depreciation—is that this gap will become bigger when the participants receiving an old sandwich are *informed* about its distant production date. In other words, we hypothesize that knowing that one's sandwich has been produced 7 days ago has a detrimental psychological effect on the gustatory experience.

Our study stands out from existing research in several ways. First, we look at the effect of freshness information on people's *experienced* consumption utility rather than only on their expectations. Second, we use a natural and realistic type of consumption—i.e., the sandwiches we use are typically consumed within the time frames of our experiment—where we make clear that the food item distributed is safe to eat, thereby isolating the sole psychological effect of freshness considerations on people's gustatory experience. Third, our *Blind* condition allows us to neatly assess the latter effect by controlling for the *objective* taste of the sandwich. Finally, contrary to existing research, our experiment is designed so as to minimize the size of eventual demand effects.

Previewing our results, first, we find that 7-day-old and 1-day-old sandwiches do not objectively differ: they do not provide a distinct consumption experience, as measured by taste ratings and WTPs, to participants who are unaware of the actual freshness of their sandwich. This remarkable stability of sandwich quality over time provides us with an ideal setting to test for the existence of psychological depreciation, as the only factor left that

could explain a difference in consumption utility between groups is the knowledge of the production date. Surprisingly, however, we do not find a psychological effect of freshness information on subjects' gustatory experience. More specifically, participants who knowingly received an old sandwich do not indicate a significantly lower experienced utility of consumption than subjects who either consciously ate a fresh sandwich or ate an old one ignoring its distant production date. Our results therefore suggest that consumers are not overly sensitive to information regarding food's freshness. That is, they appear to enjoy the food in our study equally well when it is new or old and when they do or do not know this freshness information. This finding may, for example, encourage stores to keep food for sale for longer than they might otherwise if consumers actually exhibited psychological depreciation. In the conclusion, we further elaborate on our findings.

The rest of the paper is organized as follows. The next section provides an overview of the literature. We then describe our experimental design and our main outcome variables. In section 2.4, we define our behavioral hypotheses, followed by the presentation of the results in section 2.5. We conclude in section 2.6 with a discussion of our findings.

2.2 Literature review

In the last few decades, much research, mostly in marketing and food sciences, has studied consumers' preferences for freshness. Freshness informational cues were shown to be a relevant factor in people's purchase and consumption decisions (e.g. WRAP, 2011; IFICF, 2007; Food Safety, 2009; Newsome et al., 2014; Dinnella et al., 2014; Verbeke and Ward, 2005; Harcar and Karakaya, 2005). Evidence shows that perishable food items approaching their expiry dates are perceived as less acceptable for consumption and as less safe (e.g. Sen and Block, 2008; Wansink and Wright, 2006; FSA, 2012; Tsiros and Heilman, 2005), even in cases where the latter is not justified (Ransom, 2005; NRDC, 2013). These presumptions may explain why we observe a decrease in people's expected utility over the

shelf life of food items (e.g. Kerley et al., 2008; Tsiros and Heilman, 2005; Sen and Block, 2008; De Hooge et al., 2017; Wilson et al., 2017). While this research focuses on people's *predictions* of food taste, it but does not say anything about their actual gustatory *experience*.

Only a few studies have participants tasting and rating food items. For instance, Wansink and Wright (2006) run a within-subject experiment where participants are asked to try and rate yoghurts displaying different (fake) freshness dates, such that some of them are (supposedly) expired or not. They find that people's perception of freshness and healthfulness is declining over the shelf life and beyond, with a drop at the time of expiration. Similarly, Samotyja (2015) observes a decrease in participants' acceptance of potato chips when they believe that the chips are close to their freshness date.

While the above mentioned literature focuses on people's preferences change around a product's expiration or freshness date, our interest resides more in the effects of varied *storage times* of an unexpired item on people's consumption utility. Indeed, while deciding whether to consume or not an item at home, consumers may discard foods past their expiration date, but may also throw away food that has been stored for longer durations even if an expiration date is irrelevant or has not been reached (Parizeau et al., 2015). Dinnella et al. (2014) investigated the effect of storage duration information on people's sensory evaluation of food. They presented subjects with four pre-cut and pre-washed salad samples that had been stored for 0, 3, 7 or 9 days since their conditioning. Participants were asked to try and rate them under three different conditions: with no information, with true information and with wrong information on storage time. While subjects only marginally preferred the salad freshly made to the one stored for 9 days in the no information condition, perception of freshness steadily and significantly decreased over storage time, remaining but identical between 7 and 9 days of storage in the true information

treatment. Based on the two conditions with information, they find that indication of a short (long) storage time on the packaging affect positively (negatively) how fresh the salad is perceived, independently of its true storage time. Similar effects of storage time information on experienced utility were also found for other perishable food items such as cod fish (Kole et al., 2009; Østli et al., 2012).

These findings suggest the existence of a psychological effect of freshness information on people's experienced utility. However, this evidence needs to be validated by measuring individuals' preferences through more general survey questions (i.e. asking about overall liking rates rather than freshness perception) and also through incentivized behavioral measures. Lund et al. (2005), the closest study to ours, provide some evidence in this regard. They measure subjects' expected and experienced utility in an incentivized setting, where participants express their preferences over *new* Granny Smith apples (that have been stored for 2 months) and *old* ones (that have been stored for 8 months). Depending on the treatment, participants are materially endowed with a kilo of old or new apples. They are then informed about the age of their apples and can assess them visually and through touch. Afterwards, they are randomly allocated to experimental markets of 6 people to play a Vickrey auction (Vickrey, 1961). In each out of five rounds, the subjects have to submit their willingness-to-pay WTP (willingness-to-accept WTA) to exchange their old (new) apples for a bag of new (old) apples. Then, the subjects have to knowingly try samples of apples of both ages before playing, again, five rounds in the same auction market. One out of the ten market rounds is then selected for implementation, and the buyer (i.e. the winner) pays the market price out of his earnings. Before and after having tasted the fruits, participants are also asked to state their preferences (i.e. what apples they (would) prefer to consume) and their hedonic liking (i.e. how much do (would) they like old and new apples on a scale from 1: *dislike extremely* to 9: *like extremely*). The results indicate that people

mispredict how tasty the old apples are compared to the new ones. The authors find that stated preferences and hedonic liking are much higher for the new than for the old apples *before* the tasting phase, while these ratings become similar for both apples' ages *after* tasting. Data on WTP and WTA also reveals that subjects initially value new apples more than old ones. However, the tasting phase has heterogeneous effects on people's valuations: while some consumers decide that the new apples have no value after having tasted them, some others maintain their high valuations.

While Lund et al's study reveals that participants mispredict how well food items varying in freshness will taste, it cannot cleanly identify the psychological effect of the storage duration on participants' food evaluation as the authors do not have participants evaluate food products *without awareness* of their freshness. This is not the case of our study. Our experimental setting allows us to not only compare the consumption utility of individuals who *either* received *or* did not receive freshness informational indication, but also to control for the existence of the *objective* depreciation of the sandwich's taste over time. In addition, we also make clear that the food item that we are offering is safe to eat, thereby ruling out any psychological effects that safety concerns may have on the gustatory experience. Another weakness of Lund et al's study is that subjects are put in a rather unrealistic situation where they evaluate 2- or 8-month-old apples. While supermarkets may store apples for this amount of time, households usually do not do so. Our participants face a more credible situation in which they evaluate a 1- or 7-day-old food item, a reasonable storage and consumption period for the kind of food we employ. Finally, unlike Lund et al.'s and related work, we attempt to minimize the size of eventual demand effects. We do so by providing participants with additional product information different from the production date and by running a between-subjects design where participants are unaware of the existence of other treatments (at the time of the food evaluation).

2.3 Study design

2.3.1 Design overview

We recruited students on the campus of the University of Zurich by approaching them in the corridors between 11:45 and 13:00. We asked them whether they would be interested in participating in a 20 minute study about food evaluation that would remunerate them with 10 CHF and a free sandwich. If they showed interest, and fulfilled eligibility criteria,[2] they were directed to our desk at a classroom's entrance. At the desk, participants were allocated an ID corresponding to their seat number and their self-reported hunger was measured using a one-question survey.

Once seated in the classroom, every subject was brought the first half of a sandwich that they were instructed not to start eating before the completion of Part I. After having filled out the first part of the questionnaire and eaten at least three bites of their sandwich, Part I was collected and Parts II and III were distributed one after the other. The questions in each part are discussed in more detail below. Once Part III was filled out, the experimenter brought back the Part II of the questionnaire previously completed by the same person and made her roll a die to implement an incentivized behavioral measure. Finally, participants were asked go back to the desk outside the classroom to receive their payment and, if purchased, the second half of their sandwich.

2.3.2 The treatments

The four conditions of our 2 x 2 between-subjects design are summarized in Table 1. We varied two dimensions: the freshness of the sandwich and the disclosure of its production

[2] To participate, a person should not have any allergies, intolerances or diet restrictions that would prevent them from eating the sandwich, should understand written English, and should not have previously participated in our study.

date. Participants received a sandwich that had been produced either 1 day (*Fresh*) or 7 days (*Old*) before the experiment took place, and were either informed (*Salient*) or not (*Blind*) of its production date. Importantly, all participants were explicitly told that their sandwich was safe to eat. In the rest of the paper, we use the following abbreviations for the different conditions: *BO* for *Blind-Old*, *BF* for *Blind-Fresh*, *SO* for *Salient-Old* and *SF* for *Salient-Fresh*.

Table 1: Summary of the four conditions

	Blind *(Production date* not *provided)*	**Salient** *(Production date provided)*
Old *(7-day-old sandwich)*	Blind, Old (BO)	Salient, Old (SO)
Fresh *(1-day-old sandwich)*	Blind, Fresh (BF)	Salient, Fresh (SF)

2.3.3 The sandwich

We bought packaged egg sandwiches made out of toast bread in a Swiss supermarket (a picture is provided in Appendix A, Figure A1). The packages were transported in a cool box with ice packs and were stored in a fridge at the university with the temperature continuously monitored. On the day of the experiment, we used a portable fridge if the classroom was not equipped with a refrigerator. The packages were opened right before distribution.

This type of sandwich was especially adequate for our setting in many regards. First, its compatibility with vegetarianism maximized the pool of eligible participants. Second, it consisted of two halves of 82.5 grams each. This is a relevant element as we wanted students to remain hungry after having eaten their first sandwich half since one of our measures of consumption utility is the subjects' willingness-to-pay for the second half of their sandwich. Finally, the sandwich had a long storage duration (up to eight days), and its packaging was optimized so as to well preserve its quality over time. This is important as a sandwich's

stable taste and quality provides us with an ideal environment to neatly identify the sole effect of the provision of the production date on experienced utility of consumption.

2.3.4 The survey

Overall, the survey consisted of four distinct parts: the "hunger" question, Part I, Part II and Part III. When a subject was done with one of the parts and had signaled this to the experimenter, she received the next part in exchange for the completed one.[3] Every participant's randomly allocated ID indicated her seat number and also assigned her to one of the four conditions, which varied within sessions. The subjects' ID was printed on every sheet of instructions. All instructions are available in Appendix A.

The "hunger" question

Before entering the classroom, the participants indicated their level of hunger by crossing a number between 1: *Not at all hungry,* to 9: *Extremely hungry*, on a small slip of paper. The collection of this information outside the classroom ensured that the answer was not influenced by viewing of the sandwich potentially triggering salivation or disgust.[4]

Part I: Treatment

The main purpose of Part I was to provide freshness information to participants in the *Salient* conditions. We did so by listing general information on the sandwich (i.e. the production date, the weight of the sandwich half, and its main ingredients) in the context of providing them information specific to a study on food evaluation.[5] The production date was removed from the instructions for subjects in the *Blind* conditions.

[3] This mechanism prevented subjects from changing their choices in already completed parts.

[4] Note that the traffic of sandwiches was organized in such a way that new participants would not see any sandwich before entering the classroom.

[5] To make sure that participants attentively read the general information, they were asked three control questions at the end of Part I. The experimenter systematically checked the correctness of these answers before collecting Part I and distributing Part II.

Once seated, the participants received their sandwich half on a white paper plate. They were instructed not to eat it before having rated its appearance (using a scale from 1: *Not at all appetizing* to 9: *Extremely appetizing*), its smell (1: *Smells very bad*, to 9: *Smells very good*), and having indicated how eager they were to eat the sandwich (1: *Not at all eager*, to 9: *Extremely eager*). We use the latter question as an indicative measure of subjects' expected utility of consumption.[6]

At the end of Part I, participants were instructed to eat at least three regular-sized bites of their sandwich half in order to receive the next part of the questionnaire.

Part II: Outcome variables

Part II was used to collect the two main measures of experienced utility of consumption: taste ratings and an incentivized behavioral measure, namely the willingness-to-pay for the second half of the person's sandwich.

Participants first had to report the experienced taste of their sandwich on a scale from 1: *Extremely bad*, to 9: *Extremely good*. Then, we used the Becker-DeGroot-Marschak method (Becker et al., 1964) to collect participants' willingness-to-pay (hereafter *WTP*) for the second half of their sandwich. They had to indicate, for every price increment of 0.20 CHF between 0 CHF and 2.20 CHF, whether they would buy or not buy the second half. They were informed that their answers together with their random draw of a price at the end of the experiment would determine whether they would purchase the second half. Finally, subjects had to rate how healthy and nutritious they considered their sandwich to be (from 1: *Not at all healthy and nutritious*, to 9: *Extremely healthy and nutritious*).

[6] Although it also captures a person's level of hunger, we preferred this question over asking directly for taste expectations as individuals may have reported the same answer when asked later to rate the *actual* taste of the sandwich for the sake of being consistent throughout the survey.

Part III: Manipulation check and individual characteristics

The purpose of Part III was to test if our treatments worked and to assess a wide range of individual behaviors, beliefs, opinions and preferences that we briefly define below. For a more extensive description of these variables, see Brun (2018), which examines them in more detail.

Manipulation check: A precondition for the production date to affect participants' consumption expectations and experiences is that subjects in the *Salient* conditions properly remembered the production date of their sandwich while eating it. For this reason, we asked them if they recalled whether their sandwich had been made 1 or 7 days before the day of the experiment.

Individual characteristics: Throughout Part III, we asked a wide range of questions that allow to assess individuals':

- level of familiarity with eating sandwiches,
- behaviors towards- and opinions about food waste,
- beliefs about social norms in the domain of food waste,
- health concerns,
- and risk preferences (in general, in the health/safety domain and in the food domain).

Subjects indicated whether they had already eaten this type of sandwich in the past (*Yes/No*). They were also asked if they experienced a serious case of food poisoning in the past, and in the positive case, to indicate the type of food that made them sick. This question conveys information on a person's sensitivity to food and prior aversive experiences with food. In addition, we elicited several demographic characteristics, such as gender, age, nationalities, highest achieved educational degree, and fields of study. We also asked whether they had previously heard of this study before participating and what they heard in the case of a positive answer.

When a participant was done with Part III, the experimenter brought back her previously completed Part II as well as a 12-faced die to implement the incentivized measure of WTP. The participant was instructed to roll the die to randomly draw a price p of the second sandwich half (every face of the die corresponded to a price $p \in \{0, 0.2, 0.4, \ldots, 2.2\}$ CHF), and therefore determine in combination with her answers whether she would actually buy or not at that price. If a purchase took place, the participant received the second half of her sandwich as well as a participation fee of $10 - p$ CHF. Otherwise, she simply received her full participation fee.

Participants who had completed the study but had not finished their sandwich had to either finish eating it in the classroom or to give back to us its remaining portion in order to leave. We use the weight of the unfinished first sandwich halves as a secondary behavioral measure of experienced utility, since a high (low) weight may reflect a bad (good) gustatory experience. We did not allow them to leave the classroom with their sandwich leftovers, as we would not have been able to observe if they would have thrown away the unfinished sandwich instead of eating it entirely.

We conducted 10 sessions in March and April 2018, in two classrooms located in distinct areas of the campus of the University of Zurich. Every session consisted of 22 to 29 participants. We recruited 269 participants in total but the loss of some data points[7] lead to a final dataset of 259 observations distributed across conditions as follows: 66 in the Salient-Old, 65 in the Salient-Fresh, 63 in the Blind-Old and 65 in the Blind-Fresh condition.

[7] 8 participants were mistakenly given Part III of the instructions before Part II. As the content of Part III made the problem of food waste salient, it may have influenced participants' answer to Part II, which is the reason why we dropped these observations. We also dropped one participant who did not rate the taste of the sandwich and another one who had heard too much about the study before participating.

2.4 Behavioral hypotheses

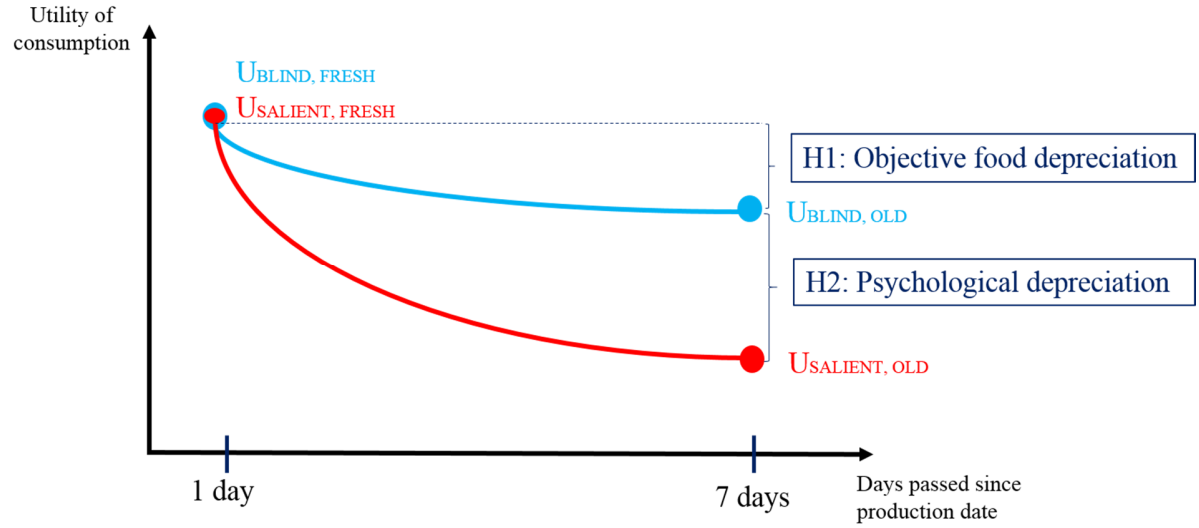
Our two behavioral hypotheses are summarized in Figure 1.

H1: The objective sandwich quality may depreciate over time

We expressly selected this type of sandwich because it is designed to persist in taste and quality over several days of storage. Thus, our goal was to have a design in which the objective quality of the sandwich is very similar after 1 or 7 days. However, we still allow for the possibility that some objective depreciation may occur and that the sandwich tastes slightly worse after 7 days of storage duration than after 1 day. Over time, the bread may for instance become drier on its outside, and softer on its inside because of the moist content of the sandwich. This assumption is tested by comparing the gustatory experience of participants in the *Blind* conditions since it is exempted from any eventual psychological influence of knowledge of the production date (see H1 in Figure 1). More formally:

$$U(Sandwich|Blind, Fresh) \geq U(Sandwich|Blind, Old)$$

Figure 1: Illustration of the behavioral hypotheses



H2: A distant production date generates psychological depreciation

Previous studies found that people's experienced utility of consumption shrinks as the storage time increases (Lund et al., 2005; Dinnella et al., 2014; Kole et al., 2009; Østli et

al., 2013). However, none of these studies can disentangle whether this trend is solely driven by the objective qualitative deterioration of the food item, or/and by the psychological effect of the distant production date on the consumption experience. We hypothesize that the simple knowledge of the sandwich's lack of freshness hinders experienced utility of consumption, phenomenon that we call *psychological depreciation* (see H2 in Figure 1). We test for its existence by doing the following difference-in-difference comparison:

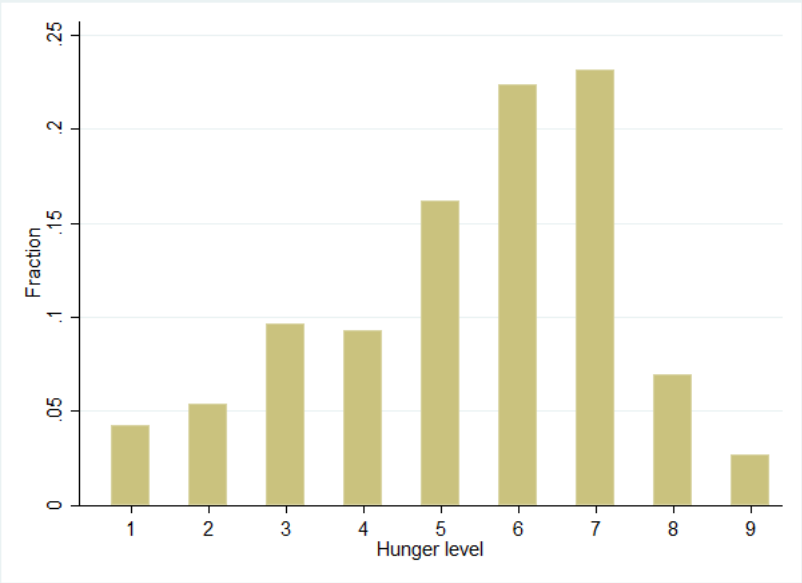
$$[U(Sandwich|Salient, Old) - U(Sandwich|Salient, Fresh)]$$
$$- [U(Sandwich|Blind, Old) - U(Sandwich|Blind, Fresh)] < 0$$

2.5 Results

Our final sample consists of 259 individuals between 18 and 35 years old, with 52% of them being women and 79% having the Swiss nationality. Only 3% of the participants had heard of our study beforehand but most did not know any critical information.[8] Importantly, *all* subjects in the *Salient* conditions were able to subsequently correctly report the production date of their sandwich, evidence that our manipulation worked. Participants' hunger level varied a lot (see its distribution in Figure 2). A two-sided t-test reveals that they were on average rather hungry ($\overline{Hungry} = 5.39$, p-value $= 0.001$). This is important as this is a prerequisite for individuals to provide a weakly positive WTP for the second sandwich half.

[8] They mostly reported having friends who recommended to participate in our study without providing them with any further information.

Figure 2: Distribution of hunger levels



Notes: Hungry=1: Not at all hungry, =9: Extremely hungry. N=259.

We define a subject's willingness-to-pay as the maximum price at which she still indicates preferring buying the second half rather than giving it up. However, we code it differently for eleven subjects who provided inconsistent answers by changing between columns more than once (see question 5 in the instructions provided in Appendix A). Answers of this kind do not provide a clear threshold price. However, they are still informative of a person's valuation of her second sandwich half. We therefore recode these observations using a formula that essentially weights the maximum price a person is willing to pay by the ratio of the sum of the prices that she agrees to pay below that level to the sum of all possible prices below that level.[9] This formula never produces a higher outcome

[9] More formally, let $k \in\{0, 0.2, 0.4 \ldots 2.2\}$ be the set of all possible prices and let x_{ik} take the value 1 if an individual i is willing to receive the second half at price k, and 0 otherwise. Define the maximum willingness-to-pay of an individual i as $k_{i,max} = \max_k k * x_{ik}$. A person i's willingness-to-pay $\widehat{wtp_i}$ is then obtained by applying the following formula: $\widehat{wtp_i} = \left(\left(k_{i,max} + 0.2\right) * \sum_{k=0}^{k_{i,max}} [x_{ik} * (k + 0.2)] / \sum_{k=0}^{k_{i,max}} (k + 0.2) \right) - 0.2$. We add 0.2 to all prices as this ensures that the formula penalizes inconsistencies at 0 (for example, a person who does not want the sandwich for free, but agrees to pay every price up to 1 CHF, is allocated a lower WTP than a person who would agree to receive the sandwich for any price from 0 (included) to 1 CHF). We but center the outcome of the formula to 0 to obtain $\widehat{wtp_i}$.

than the participant's maximum reported price, sensibly responds to different amounts chosen below the maximum one, and delivers the maximum price for consistent people (i.e those who switch columns only once).

Out of 259 participants, 60[10] did not finish their first half, and 37[11] did not even want the second half of the sandwich for free (coded as WTP<0 in Figure 3b). Overall, the sandwiches were not especially well or badly evaluated. The average WTP for the second sandwich half reached 0.78 CHF (among the 222 participants who indicated a weakly positive WTP). The average taste and smell ratings of 5.35 and 5.92, respectively, were both significantly higher than 5 and the sandwiches were judged slightly less appetizing than 5 with 4.67 (the p-values of two-sided t-tests are all <0.001). Individuals experienced very heterogenous levels of consumption utility as indicated by the fair distributions of taste ratings, WTP and weight of unfinished first halves (see Figures A2a, A2b and A2c in Appendix A, respectively). This high variance may not only ensue from our treatments but also from the wide range of hunger levels among participants (see Figure 2). General summary statistics of the relevant variables of our study are available in Table A1 (Appendix A). For an analysis of the data collected on participants' personality traits and reported food waste behaviors, please refer to Brun (2018).

Throughout the paper, we do not focus our main analyses on the weights of the unfinished first sandwich halves, as many reasons others than a bad gustatory experience have led to these leftovers.[12] In addition, subjects were eating the sandwich while answering an extensive survey on opinions, beliefs and behaviors towards food waste. This may have affected their decision to finish or not finish the sandwich.

[10] These 60 observations are distributed across conditions as follows: N_{BO}: 11, N_{BF}: 19, N_{SO}: 16, and N_{SF}: 14.
[11] These 37 observations are distributed across conditions as follows: N_{BO}: 6, N_{BF}: 11, N_{SO}: 12, and N_{SF}: 8.
[12] Subjects typically reported having to leave to attend a lecture, having just had lunch or being willing to have a proper lunch.

2.5.1 Objective difference between a 7-day-old and a 1-day-old sandwich

In this section, we investigate whether the experienced consumption utility *objectively* (i.e. without any psychological influence of freshness information) differs between an old and a fresh sandwich. For this purpose, we focus exclusively on the Blind conditions. If the sandwich's quality decreases over time, subjects in the Blind-Old condition would indicate lower taste ratings and would be willing to buy the second half of their sandwich at a lower price than individuals in the Blind-Fresh group. Also, if a difference is noticeable between a 7-day-old and a 1-day-old sandwich, subjects would provide an accurate guess of their sandwich's freshness and the quantity of unfinished first halves would be lower in the Blind-Fresh than in the Blind-Old condition.

Figures 3a, 3b and 3c plot the cumulative distribution function of participants' taste ratings, WTP, and of their guess about their sandwich's freshness, respectively. At first sight, there is no obvious visual difference of the distributions between the Blind-Old and the Blind-Fresh conditions, independently of the variable considered.

We perform two-sided Wilcoxon rank-sum tests and t-tests to test for differences in distributions and in means, respectively. We do not find any significant differences in the taste ratings' distributions and in the average taste ratings between the Blind-Old and the Blind-Fresh conditions ($\overline{Taste}_{BO}$=5.49, $\overline{Taste}_{BF}$=5.45, p-value>0.1 for both tests). The same conclusion holds when examining the difference in WTPs ($\overline{WTP}_{BO}$=0.74, $\overline{WTP}_{BF}$=0.86, p-value>0.1 for both tests)[13] and in guesses about when the sandwich was made ($\overline{Guess}_{BO}$=4.08, $\overline{Guess}_{BF}$=4.15, p-value>0.1 for both tests).[14] The distribution and the mean

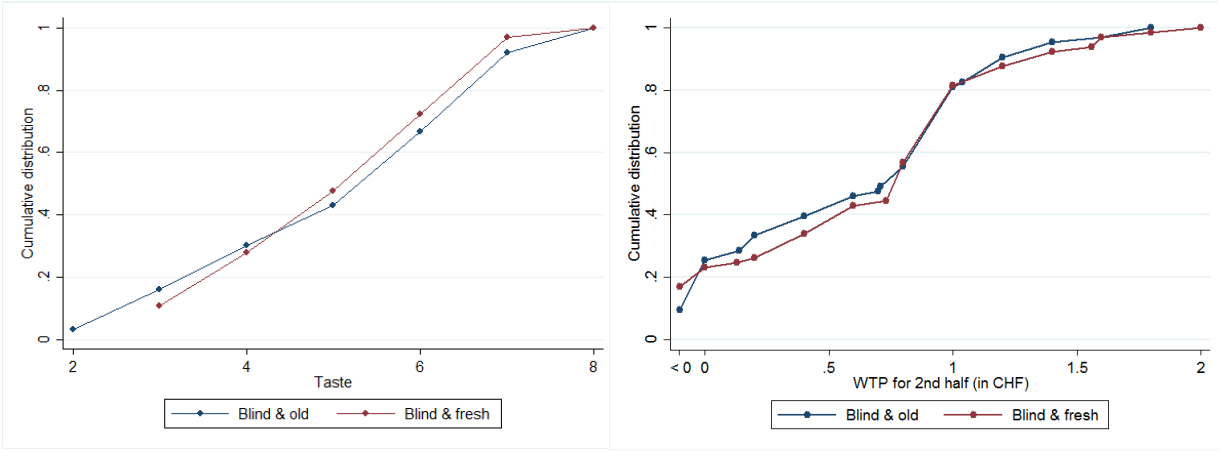
[13] Note that throughout the paper, tests to compare the *means* of WTPs across groups are performed on a reduced sample (N_{BF}=54, N_{BO}=57) since we remove observations with a negative WTP. This concerns 11 (6) observations in the Blind-Fresh (Blind-Old) condition, respectively. However, as rank-sum tests treat the data as ordinal, we perform them on the entire sample.

[14] Subjects who are more familiar with eating sandwiches provided more accurate guesses. However, they neither experienced a different taste nor indicated a different WTP than subjects who do not regularly eat sandwiches.

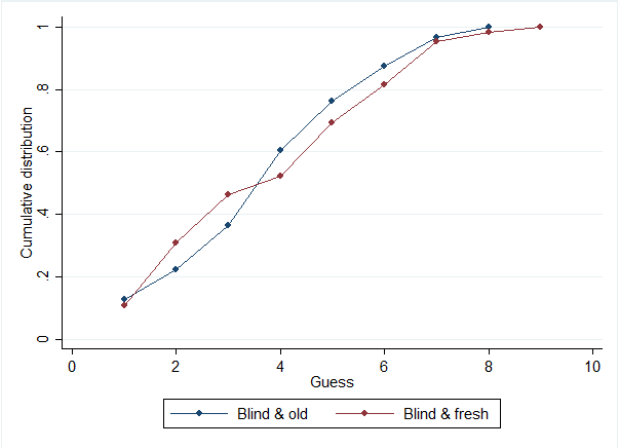
of the weight of unfinished first halves are also not significantly different between the two groups ($\overline{Weight}_{BO}$=10.4, $\overline{Weight}_{BF}$=16.5, p-values>0.1 for both tests).

The inability of subjects to correctly guess how fresh their sandwich is may not be only explained by the absence of gustatory difference, but also by the similar appearance and smell of both types of sandwiches (see Table A2, Appendix A, for the tests of means and distributions of smell and appearance ratings). Interestingly, participants tend to believe that their sandwich is rather fresh, independently of the condition in which they are ($\overline{Guess}_{BO}$ and $\overline{Guess}_{BF}$ are both significantly lower than 5, p-values<0.05 of two-sided t-tests), although they are aware of their equal chance to have been allocated an old or a fresh sandwich.

Figure 3: Cumulative distribution function of taste ratings (a), willingness-to-pay for the second half (b) and of participants' guess about the freshness of their sandwich (c) in the blind conditions



(a) **(b)**



(c)

Notes: Taste=1: Extremely bad taste, =9: Extremely good taste; WTP in [0; 2.2] CHF, WTP=-0.1 if a person does not even want the second half for free; Guess=1: Was definitely made 1 day ago, =5: I can't really tell, =9: Was definitely made 7 days ago. N_{BO}*=63 and* N_{BF}*=65.*

To summarize, we do not find evidence for a clear difference in the quality of a 7-day-old and of a 1-day-old sandwiches. These two types of sandwiches do not provide distinct gustatory, visual or olfactory experiences, and are similarly valued financially. Subjects are also unable to correctly guess the freshness level of their sandwich. This absence of actual food quality depreciation over time is ideal in our setting as it allows us to rule out the effect of quality deterioration on our measures of experienced utility. In other words, any evidence for unequal utility of consumption between old and fresh sandwiches must be solely triggered by the psychological effect of freshness information. We investigate this in the following section.

2.5.2 Psychological effect of the date on experienced utility

In this section, we investigate whether freshness information has a direct (psychological) effect on individuals' gustatory experience. More specifically, we test if the simple fact of being aware that one's sandwich has been produced 7 days ago negatively affects one's utility of consumption, a phenomenon that we call *psychological depreciation*. The absence of quality deterioration found in the previous section implies that the

differences in experienced utility that we would find between the Blind-Old and the Salient-Old conditions, and between the Salient-Fresh and the Salient-Old conditions, would solely be driven by the knowledge of the distant production date. It thus makes sense to start this section by comparing the average and distribution of experienced utility between the relevant conditions, and refine our investigations by running regression analyses.

Mean and distributions comparisons between treatment groups

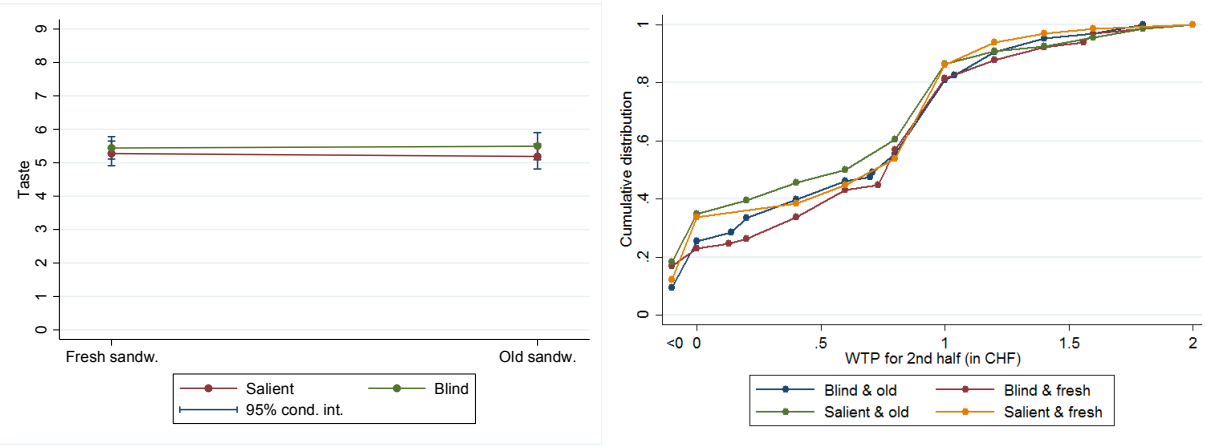
Figures 4a and 4b plot the average taste ratings and the cumulative distribution of WTPs in each of the four conditions, respectively. We do not primarily consider *average* WTPs since this would imply ignoring individuals who do not even want their second half for free, as we do not know what financial compensation they would require to receive it.[15] There is visually and statistically no significant differences between the Blind-Old and the Salient-Old conditions either when performing two-sided t-tests of the means of taste ratings ($\overline{Taste}_{BO}$=5.49, $\overline{Taste}_{SO}$=5.19, p-value=0.283) or two-sided Wilcoxon rank-sum tests of taste ratings and WTPs (the p-value is above 0.1 in both cases). We draw the same conclusion when considering the Salient-Fresh and Salient-Old conditions ($\overline{Taste}_{SF}$=5.28, $\overline{Taste}_{SO}$=5.19, p-value of the two-sided t-test is 0.744 and the p-values of two-sided Wilcoxon rank-sum tests on taste ratings and on WTPs are above 0.1).

In a second step, we examine the weight of the remaining first sandwich halves in the four conditions. The results are qualitatively aligned with our hypothesis: the quantity of leftovers is higher when subjects are aware of the oldness of their sandwich (Salient-Old) than when they are not (Blind-Old) or than when they knowingly received a fresh sandwich (Salient-Fresh) ($\overline{Weight}_{BO}$=10.43, $\overline{Weight}_{BF}$=16.46, $\overline{Weight}_{SO}$=13.15, and $\overline{Weight}_{SF}$=10.58). However, none of the relevant mean and distribution differences are

[15] However, we provide the graphical representation of the average WTP in each condition (calculated on the reduced sample of weakly positive WTPs) in Figure A3 in Appendix A.

significant, even at a 10% level. The large confidence intervals remind us that the weight of sandwiches' leftovers is a noisy measure of consumption utility, and that our main conclusions should not mainly rely on this metric.

Figure 4: Average taste ratings and cumulative distributions of willingness-to-pay in each of the four conditions



(a) **(b)**

Notes: (a): Taste=1: Extremely bad taste, =9: Extremely good taste. (b): WTP in [-0.1; 2.2] CHF. N=259 with N_{BO}=65, N_{BF}=63, N_{SO}=65 and N_{SF}=66.

This preliminary set of results points towards the absence of psychological depreciation: participants who are informed about the distant production date of their sandwich do not experience a significantly lower gustatory experience than subjects who either are not aware of the lack of freshness of their sandwich or know that their sandwich was produced on the previous day. In the following paragraphs, we test for the existence of psychological depreciation by running several difference-in-differences estimations where we account for potential treatment randomization failures.

Regression analyses

Table 2 displays the results of linear (columns 1 to 3) and of ordered Probit (columns 4 to 6) regressions of the taste ratings on the treatment and on control variables. Columns

1 and 4 indicate that none of our treatment has a significant effect on the participants' taste ratings. In particular, there is no evidence for the existence of psychological depreciation ($\hat{\beta}_{Salient*Old}$ is not significant). These findings are not surprising considering the very low explanatory power of these models (R^2=0.007 in column 1, and pseudo R^2=0.002 in column 4).[16] Controlling for individual characteristics (columns 2, 3, 5 and 6) and sessions' fixed effects (columns 3 and 6) does not impact any of the treatment effects. As a robustness check, we recode the taste ratings such that they only consist of three categories[17] and find similar results (see Table A3 in Appendix A).

Out of 259 participants (with N_{BO}=65, N_{BF}=63, N_{SO}=65 and N_{SF}=66), 37 of them did not even want the second half of their sandwich for free, and 39 subjects indicated a willingness-to-pay of 0. To account for these two types of preferences, we implement a hurdle model which consists of a Probit model (to explain the acceptance or refusal to pay a (weakly) positive price for the second sandwich half), and of a linear model truncated at 0 (to predict the price a person is willing to pay *conditional* on having accepted to pay a weakly positive price). Results are reported in Table 3. As for the taste ratings, variations in our treatments do not induce substantial changes in subjects' willingness-to-pay (pseudo R^2=0.012 in column 1) which is further reflected by the absence of significant treatment effects in columns 1 and 2 ($\hat{\beta}_{Salient*Old}$ is not significant). When adding control variables to the model (columns 3 to 6), the variables *Old* and *Salient*Old* seem to significantly explain the subjects' decision to pay or not a weakly positive price for their second sandwich half ($\hat{\beta}_{Old}$

[16] R^2 statistics are not available for ordered Probit and Tobit models but have an equivalent metric called the McFadden pseudo R^2. This measure is obtained by the following formula: $R^2_{McFadden} = 1 - \frac{\log(L_{Full})}{\log(L_{Int.})}$, where L_{Full} ($L_{Int.}$) denotes the maximized likelihood value of the full model (of the model that only contains an intercept), respectively. Therefore, a value close to 1 (0) indicates a high (low) explanatory power of the model, respectively.

[17] We replaced *Taste* by *Taste cat.* where *Taste cat.*=1 if the person indicated a low taste rating (i.e. *Taste* <5), =2 if the person's gustatory experience was neither bad nor good (i.e. *Taste* =5), and =3 if the subject indicated a high taste rating (i.e. *Taste*>5).

and $\hat{\beta}_{Salient*Old}$ are significant at the 10% and 5%-level, respectively, in both columns 3 and 5), but the different treatments can still not explain the financial amount individuals are willing to pay across conditions[18] ($\hat{\beta}_{Old}$, $\hat{\beta}_{Salient}$, and $\hat{\beta}_{Salient*Old}$ are not significant in columns 4 and 6). While these results seem to show some support for the existence of psychological depreciation, a deeper inspection of the data casts some doubts on this conclusion. First, a simple visual analysis of the average WTPs in each condition (see Figure A3 in Appendix A) does not corroborate the hypothesized relationship between treatments. Also, the coefficients of our three treatment variables are jointly only weakly significant (the p-value of a Chi-squared test on the Probit models in columns 1, 3 and 5 equals 0.464, 0.095 and 0.059, respectively). Finally, running linear- and ordered Probit regressions on the whole sample by specifying WTP as a categorical variable (with *WTP cat.*=1 if *WTP*<0, *WTP cat.*=2 if *WTP*=0, and *WTP cat.*=3 if *WTP*>0) do not indicate any significant differences in WTPs across the different conditions (see Table A5, Appendix A).

[18] Note that linear as well as Tobit regressions implemented on a reduced sample without negative willingness-to-pay values do also not produce any significant treatment effects (see Table A4 in Appendix A).

Table 2: Relationship between the taste ratings and the treatments

Dependent variable:	**Taste**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Old	0.0459	0.104	0.113	0.0440	0.0837	0.0925
	(0.17)	(0.40)	(0.43)	(0.24)	(0.47)	(0.52)
Salient	-0.169	-0.182	-0.200	-0.110	-0.127	-0.142
	(-0.66)	(-0.71)	(-0.77)	(-0.65)	(-0.73)	(-0.81)
Salient*Old	-0.133	-0.249	-0.228	-0.108	-0.193	-0.181
	(-0.35)	(-0.67)	(-0.60)	(-0.42)	(-0.75)	(-0.70)
Already eat.		0.462**	0.437**		0.332**	0.318**
		(2.36)	(2.20)		(2.47)	(2.36)
Hungry		0.108*	0.127**		0.0773**	0.0919**
		(1.95)	(2.30)		(2.03)	(2.43)
Food pois.		-0.0275	0.00137		-0.0245	-0.0124
		(-0.14)	(0.01)		(-0.18)	(-0.09)
Female		0.0473	0.0599		0.0662	0.0822
		(0.25)	(0.31)		(0.52)	(0.63)
Age		-0.0633**	-0.0530*		-0.0436**	-0.0370*
		(-2.03)	(-1.68)		(-2.03)	(-1.71)
Constant	5.446***	5.986***	5.791***			
	(31.54)	(7.63)	(6.85)			
N	259	259	259	259	259	259
R^2	0.00656	0.0616	0.0868	0.00195	0.0184	0.0257
Session FE	No	No	Yes	No	No	Yes

*Notes: Columns 1 to 3: Coefficient estimates of linear regression models. Columns 4 to 6: Estimated marginal effects of ordered Probit models. Dependent variable: Taste=1: Extremely bad taste, =9: Extremely good taste. Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. Estimated cuts for ordered Probit models are omitted from the table, McFadden pseudo R^2 are provided in columns 4 to 6, robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table 3: Relationship between the willingness-to-pay and the treatments using a hurdle model

Dependent variable:	WTP weakly pos. (1/0)	WTP	WTP weakly pos. (1/0)	WTP	WTP weakly pos. (1/0)	WTP
	(1)	(2)	(3)	(4)	(5)	(6)
Old	0.352	-0.0431	0.546*	-0.0426	0.557*	-0.0497
	(1.23)	(-0.50)	(1.92)	(-0.50)	(1.95)	(-0.60)
Salient	0.203	0.0419	0.245	0.0234	0.245	0.0119
	(0.74)	(0.55)	(0.85)	(0.31)	(0.86)	(0.16)
Salient*Old	-0.603	0.00877	-0.897**	0.00911	-0.976**	0.0511
	(-1.53)	(0.07)	(-2.29)	(0.08)	(-2.46)	(0.44)
Already eat.			0.419**	0.113*	0.503**	0.122**
			(2.02)	(1.83)	(2.34)	(2.14)
Hungry			0.152***	0.0208	0.135***	0.0202
			(3.07)	(1.39)	(2.71)	(1.34)
Food pois.			-0.00249	0.0658	0.115	0.0802
			(-0.01)	(1.08)	(0.53)	(1.31)
Female			-0.427**	-0.0578	-0.539**	-0.0617
			(-2.12)	(-1.00)	(-2.57)	(-1.08)
Age			-0.0435	-0.00257	-0.0384	-0.00285
			(-1.46)	(-0.28)	(-1.20)	(-0.31)
Constant	0.957***	0.923***	1.138	0.809***	0.881	0.800***
	(5.18)	(15.74)	(1.55)	(3.61)	(1.05)	(3.16)
Sigma		0.391***		0.383***		0.369***
		(14.64)		(14.40)		(14.27)
N	259	183	259	183	259	183
Pseudo R^2	0.0124		0.101		0.150	
Session FE	No	No	No	No	Yes	Yes

*Notes: Columns 1, 3 and 5: Estimated marginal effects of Probit models. Dependent variable: WTP weakly pos.=1 if WTP>=0, and WTP weakly pos.=0 if WTP<0. Columns 2, 4 and 6: Estimated marginal effects of a truncated regression at 0. Dependent variable: WTP in [0; 2.2] CHF. Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. R^2 for the Probit models are McFadden pseudo R^2. R^2 are not available for truncated models. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

We turn now to the analyses of our secondary measure of consumption utility, namely the weight of the unfinished first sandwich halves. Given the low number of observations with a positive weight (only 60 cases), we assess the effects of our treatments on the likelihood that a person does not produce food waste. For this purpose, we generate a binary variable that equals 1 if a person has finished her sandwich and 0 otherwise. The absence of robust treatment effects on taste ratings and WTP is partly supported by regression results regarding the weight of the leftovers (see the estimated Probit models in Table 4). The coefficient estimates for *Salient* are not significant (columns 1 to 3). The same holds true regarding *Old* and *Salient*Old* (column 1, p-value>0.1) except when control variables are added to the model (columns 2 and 3). While linear probability models produce similar results (see Table A6, Appendix A), this apparent evidence for psychological depreciation should be interpreted with caution. First, and similarly to the analysis of WTPs, this phenomenon does not show up when comparing the average probability of finishing the first sandwich half in each of the four treatment groups (see Figure A4 in Appendix A for a graphical representation). This is not surprising given the very similar cumulative distribution of the leftovers' weight across conditions (see Figure 5). Also, we cannot confidently reject a joint test of the absence of the treatments effects when considering the estimated Probit model (the p-value of a Chi-squared test is 0.14) or linear probability model (an F-test produces a p-value of 0.09). Further analysis suggests that the significance and positive sign of $\hat{\beta}_{Old}$ in columns 2 and 3 from Table 4 is driven by unobserved individual characteristics that make subjects in the Salient-Old condition finish their sandwich more than what our control variables predict.[19] This in turn has direct consequences on the estimated magnitude and significance of $\hat{\beta}_{Salient-Old}$.

[19] Our control variables predict that 73% of the subjects in the Blind-Old condition should finish their sandwich, while 83% actually do so. This difference is much smaller when comparing the other conditions:

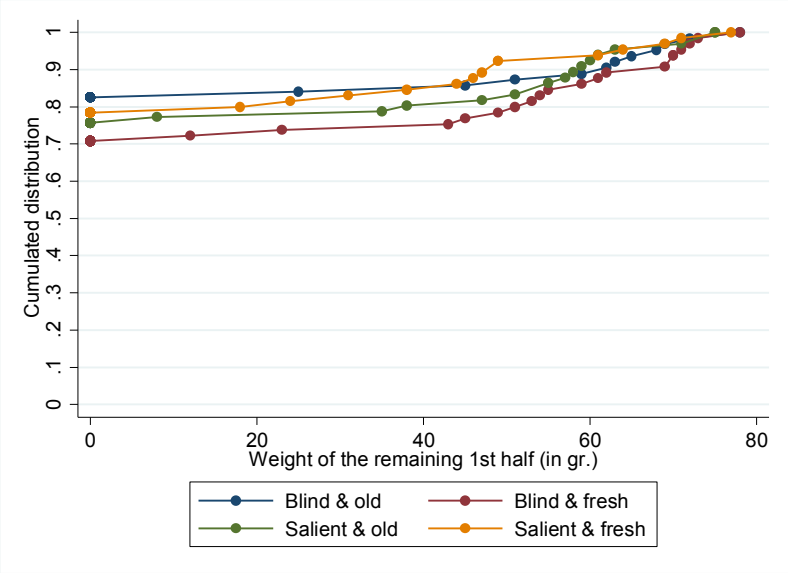
Table 4: Relationship between the decision to finish entirely the first half of the sandwich and the treatments

Dependent variable:	First half entirely eaten		
	(1)	**(2)**	**(3)**
Old	0.118	0.148**	0.152**
	(1.58)	(2.21)	(2.29)
Salient	0.0728	0.0679	0.0703
	(1.01)	(1.03)	(1.05)
Salient*Old	-0.145	-0.193**	-0.211**
	(-1.39)	(-2.09)	(-2.28)
Already eat.		0.143***	0.159***
		(3.18)	(3.70)
Hungry		0.0559***	0.0528***
		(5.00)	(4.72)
Food pois.		-0.0187	-0.0190
		(-0.37)	(-0.37)
Female		-0.223***	-0.243***
		(-5.02)	(-5.23)
Age		-0.00796	-0.0101
		(-1.05)	(-1.30)
N	259	259	259
Pseudo R^2	0.00942	0.204	0.226
Session FE	No	No	Yes

*Notes: Estimated marginal effects of Probit models. Dependent variable: First half is entirely eaten in {0;1}. Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. McFadden pseudo R^2 are provided, robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

($\bar{\hat{p}}_{SO} = 0.79, \bar{p}_{SO} = 0.76$; $\bar{\hat{p}}_{SF} = 0.78, \bar{p}_{SF} = 0.78$; $\bar{\hat{p}}_{BF} = 0.77, \bar{p}_{BF} = 0.71$, where $\bar{\hat{p}}$ ($\bar{p}$) is the predicted (actual) average probability of finishing the first sandwich half, respectively).

Figure 5: Cumulative distribution of the weight of unfinished first sandwich halves by condition



Notes: N=259 with N_{BO}*=65,* N_{BF}*=63,* N_{SO}*=65 and* N_{SF}*=66.*

To sum, we have seen that our treatments are rather bad predictors of variations in our primary (taste ratings and WTP) and secondary (weight of unfinished first halves) measures of subjects' consumption utility. The participants' gustatory experience does not seem to be significantly affected by freshness information. In particular, mean and distribution comparisons as well as regression analyses did not deliver robust evidence for the existence of psychological depreciation.

2.6 Conclusion

Lately, there has been a growing discussion around the issue of food waste and the urgent need to address it. As a large share of food waste arises at the consumer level, it is particularly important to understand the factors that influence households' decisions to discard edible food. We conduct a lab-in-the-field experiment to better understand the potential psychological mechanism whereby temporal cues influence perceptions of foods' taste and freshness, leading individuals to dispose of edible food.

More precisely, we investigate the effect of disclosing freshness informational indications of a packaged sandwich on people's experienced consumption utility. We do so by distributing halves of sandwiches varying in freshness levels (produced 1 or 7 days before the experiment) to students on a university campus over the lunch break and by revealing or not revealing the production date and duration of storage. We hypothesized that participants informed about the lack of freshness of their sandwich would enjoy their sandwich to a lesser extent than individuals who were either unaware of when their sandwich was produced or who knowingly received a fresh sandwich. In other words, we expect the disclosure of a distant production date to generate what we call the *psychological depreciation* of food.

We first show that participants who are *unaware* of the production date of their sandwich do not notice any difference in taste, smell or appearance between old and fresh sandwiches, and do not financially value them distinctively. This finding provides an ideal setting to test for the psychological effect of a long storage date on experienced utility, as the sandwich's actual quality does not appear to be influenced by time. However, we do not find robust evidence that the disclosure of the production date significantly affect either taste ratings, individuals' willingness-to-pay for a 2nd sandwich half or the weight of sandwich's leftovers. Hence, we find no evidence supporting our hypothesis of psychological depreciation.

These results may be explained by the absence of underlying psychological mechanism, namely the effect of freshness information on *expected* consumption utility that we initially aimed at testing in a follow-up study. Indeed, we find suggestive evidence for the lack of influence of freshness information on subjects' expectations, i.e. individuals who knowingly received an old sandwich were not less eager to eat their sandwich than subjects who knew that their sandwich was fresh (see Table A7 in Appendix A).

A cautious interpretation of our findings is that, by removing the likelihood that extended duration impacts food safety—recall that our subjects were explicitly told that their sandwiches were safe to eat and had not exceeded any maximum storage dates—we removed the main impediment to eating old food. That is, while previous studies demonstrate that people are more reluctant to eat food close to or past the expiration date, these studies typically confound the possibility that the food may be unsafe. In our study, on the other hand, when this is not a factor, we find virtually no difference in subjects' evaluations between sandwiches that are fairly fresh and considerably older. Thus, psychological depreciation alone—the idea that old food is simply perceived as tasting less good—may not be an important factor underlying food waste.

Of course, there remain other explanations. One explanation may come from the previous finding that incentives reshape people's beliefs about what the outcome entails, and in turn influence their decisions (Ambühl, 2017). In our experiment, the sandwich is part of the reward for participation. Individuals may thus have convinced themselves of the decent taste of their sandwich. The fact that participants in the Blind-Old condition believe on average that their sandwich is rather fresh actually supports this mechanism. Finally, an alternative reason for the lack of psychological depreciation may be that, in our setting, participants are forced to evaluate the smell and appearance of their sandwich. This first sensory evaluation may have raised their presumptions about the taste of old food and thus, may have made them more objective. In addition, the framing of our study's purpose—evaluating food—may have encouraged participants to rate the sandwich as an expert would, namely, based on objective sensory factors and voluntarily abstracting from any psychologically influencing cue.

To the best of our knowledge, we are the first paper that studies the causal connection between information on the food's production date and gustatory experiences in a real-

choice setting, and with the purpose to better understand food waste behaviors. Our findings not only complement the existing literature but also have direct implications for governments and food retailers. First, evidence shows an important drop in subjects' experienced consumption utility when the label (falsely) suggests that the item is expired or almost expired (e.g. Wansink and Wright, 2006; Samotyja, 2015). However, it remains unclear whether the negative effect of post-expiration on consumption utility is solely driven by people's safety concerns or also by their freshness considerations. Our setting allows us to rule out the first (safety) channel, as our participants are explicitly told that their sandwich is safe to eat. Therefore, our results indicate that when individuals consider eating or not an old food item, health prevails over perceived depreciation of quality. This finding is interesting from a policy perspective: Brun (2018) shows that individuals hold rather conservative beliefs when it comes to food safety. Correcting people's beliefs and misperceptions about food safety, e.g. through education, may thus be an effective way to reduce people's excessive tendency to discard old food. Also, policymakers may reconsider whether to impose the indication of the production instead of the expiry date on perishable food items since individuals' expected and experienced consumption utility seem to be less affected by the former. Second, our results may also be of interest for actors in the food industry. As individuals in our experiment neither seemed to *expect* a difference in taste nor *experienced* different taste for an old and a fresh food item, retailers may reconsider selling instead of discarding food items that have passed their freshness date. Unlike expired products, goods that have passed their *best before* date can still be sold[20] and may be equally valued by consumers. Nevertheless, retailers will often remove such items from shelves under the perception that consumers will not buy them. Results like ours suggest that such food should not be undervalued. Making the safety of these items public knowledge and

[20] This is for instance the case for Switzerland, Germany and the USA.

eventually offering a price reduction in cases where the quality actually depreciates would prevent food waste and monetary losses from discarding usable items. Retailers could also invite their clients to help them reduce food waste by redirecting their purchases towards old food products. This measure would not only make them save money, but additionally, they would potentially benefit from a better image towards their clients as evidence suggests (Theotokis et al., 2012). Note that the success of companies such as *Ässbar* in Switzerland or *Sirplus* in Germany, specialized in the resale of discounted old (and safe) food items, proves the success of this marketing strategy.

Chapter 3 – Exploring Attitudes Towards Food Waste[21]

3.1 Introduction

In the last few years, the issue of food waste has attracted growing attention from governments, NGOs, businesses, and scholars from various disciplines. This is not surprising: worldwide, 1.3 billions of tons of food is wasted every year, which represents approximately one third of all food produced for human consumption (FAO, 2013) and a yearly financial loss estimated at USD 750 billion (FAO, 2013). The consequences of food waste are not only financial, but also societal and environmental. Reducing the volume of global food waste will be key to ensuring food security for future worldwide populations (Godfray et al., 2010; Parfitt et al., 2010) and to slow down the growing demand for agricultural land that puts pressure on already shrinking forests (FAO, 2013). Food waste also heavily contributes to the release of greenhouse gas emissions (e.g. Venkat, 2011; WRAP, 2008a, 2009a) and to water wastage (e.g. Lundqvist et al., 2008). While food waste takes place along the entire food supply chain, half of it occurs at the household level (Gooch et al., 2010; Stenmarck et al., 2016). For instance, in the UK, the amount of food wasted by individuals corresponds to emitting 330 kg of CO_2 per year (i.e. equivalent to one third of the yearly electricity consumption per person, WRAP, 2011a, 2011b), accounts for about 243 liters of water per person wasted every day (WRAP, 2011a), and costs yearly around £700 to an English household with children (WRAP, 2017). Altogether, these facts emphasize the necessity to figure out the factors leading individuals to discard edible food.

Although it is urgent to better understand individuals' perceptions and behaviors related to food waste, there is for now still little research on this topic. As consumers play an important role in global food waste generation, the literature has principally investigated

[21] Please cite as Fanny Brun (2018) "Exploring Attitudes Towards Food Waste," Working Paper.

individuals' food waste perceptions and practices (see Hebrok and Boks, 2017, and Schanes et al., 2018 for literature reviews). The causes of food waste at the consumer level have been shown to be complex and multifaceted. In the current study, we complement the existing literature by looking at potential psychological drivers of food waste behaviors by using part of the survey and behavioral data gathered by Brun and Weber (2018) in their lab-in-the field experiment. We investigate to what extent risk preferences, descriptive norms and individuals' attitudes towards food waste generation are relevant predictors of self-assessed food waste behaviors, and which of these personal characteristics forms the best predictor. We also test the validity of our measure of reported food waste practices by comparing it to subjects' *actual* behavior towards food waste in the experiment previously mentioned.

In the following, we briefly review existing research that sheds light on the main determinants of individuals' food disposal behaviors, with a focus on the role played by food risks perceptions, norms and attitudes. We then explain in more details what our contribution is.

3.1.1 Previous findings

Several factors were found to be directly or indirectly related to consumers' food waste practices. First, many studies investigated what are the socio-demographic characteristics of the biggest food waste producers. This part of the literature is relevant as it provides information on *who* policies or initiatives should primarily target. It was typically found that younger people (e.g. Brook Lyndhurst, 2007; Visschers et al., 2016; Quested et al., 2013), women (Visschers et al., 2016, Secondi et al., 2015), more educated individuals (WRAP, 2008b; Secondi et al., 2015), and those with a higher disposable income (WRAP, 2008b) generate more food waste.[22] In the current study, we also look at what demographic

[22] While we are interested in food waste practices at the individual level, note that another part of the literature looked at what households' characteristics are associated with the tendency to throw away edible food. It was

traits are related to the tendency to waste food, although the sample in the dataset is limited in this regard. Our main focus is on the analysis of more fundamental drivers of food waste behaviors such as preferences, beliefs and attitudes.

Another facet of the literature investigates *why* individuals discard edible food. Knowing the causes of food waste is important as it provides key information to any institution aiming to design effective measures against food waste. Many studies emphasized what *behaviors* lead households' purchased food to become spoiled. Several domestic food-related practices and habits, in particular, food management skills, were found to be at the origin of households' food waste (see e.g. WRAP, 2009a; Brook Lyndhurst, 2007; Evans, 2011, 2012; Exodus, 2007; Stefan et al., 2013; Graham-Rowe et al., 2014; Porpino et al., 2016). Individuals were also typically shown to cook, prepare and serve too much (WRAP, 2009a) as a way to reinforce their image of "good food provider" towards their family members or guests (e.g. Graham-Rowe et al., 2014; Porpino et al., 2016; Visschers et al., 2016). More importantly for the current study, research also looked at what *psychological factors* are related to individuals' disposal of edible food. In particular, financial motivations, food risks, social norms, and concerns about food waste were shown to be relevant determinants of food waste behaviors.

Earlier work found that people's motivation to minimize the amount of food they throw away is mainly financial rather than environmental or social, as discarding food is perceived as a waste of money (e.g. Quested et al., 2013; Neff et al., 2015; Graham-Rowe et al., 2014) and as individuals are not aware of the various detrimental impacts of food waste (Brook Lyndhurst, 2007; Quested et al., 2011). Financial motivations were further shown to be positively associated with lower amounts of reported food discarded (Abeliotis et al.,

found that families with more than two adults or with children (Koivupuro et al., 2012; Parizeau et al., 2015; Williams et al., 2012), or with a larger food budget (Parizeau et al., 2015) contribute more to food waste.

2014). In the current paper, we focus on different motives associated with food waste behaviors, namely, on individuals' fear of foodborne illness, on their willingness to behave consistently with social norms or simply on people's intrinsic desire to reduce the amount of food that they discard. In the following, we provide an overview of what was previously found in the literature in this regard.

Food risk: Food safety concerns are an important determinant of households' food disposal (Neff et al., 2015; Evans, 2011; Graham-Rowe et al., 2014; Visschers et al., 2016). This is particularly true for individuals with a strong fear of foodborne illness (Exodus, 2007). Research showed that this fear enters directly in conflict with a desire to avoid throwing away food (Watson and Meah, 2013); however, consumers still prefer discarding suspicious food rather than taking the risk to be poisoned (Graham-Rowe et al., 2014).

Social norms about food waste: subjective norms (i.e. "what I think people think of my behavior") increase individuals' intention to reduce food waste but do not affect their behaviors per se (Graham-Rowe et al., 2015; Visschers et al., 2016; Stancu et al., 2016). More interestingly for the current study, descriptive norms (i.e. "what I think most people do") do not significantly predict individuals' intention to reduce the amount of edible food discarded (Graham-Rowe et al., 2015).

Attitudes towards food waste: Individuals' own opinion about the importance of reducing food waste may also affect the way they manage food items. Indeed, despite being a widespread practice, individuals consider wasting food as inappropriate behavior (Porpino et al., 2015) and feel guilty about it (Brook Lyndhurst, 2007; Stefan et al., 2013; Quested et al., 2013; Parizeau et al., 2015). Also, individuals holding personal norms (i.e. having own moral standards (Schwartz, 1977)) that oppose food waste or who are concerned with this issue (Visschers et al., 2016; Principato et al., 2015) report wasting less food.

3.1.2 Contribution

The current study contributes to this literature in several ways. First, our work provides an additional rigorous test of existing findings. Previous research generally approximated individuals' overall food waste behaviors based on a *single* survey question about subjective estimations of the frequency or amount of food wasted. The aggregated nature of our measure of food waste behaviors allows a reduction of the bias induced by potential measurement errors. Moreover, unlike earlier work, we look at the behavioral validity of our construct of reported food waste practices by testing for its capacity to predict participants' *actual* behaviors towards food disposal in the context of an experiment.

Second, to the best of our knowledge, no research so far has looked at the link between descriptive norms and reported behaviors related to food waste. As mentioned earlier, Graham-Rowe et al. (2015) only tested for the effect of this variable on people's *intentions* to reduce the volume of edible food that they discard.

More generally, we are also the first study that aims at ranking risk preferences, descriptive norms and attitudes towards food waste reduction according to how well each of these variables predicts reported food waste practices. This information is crucial to any institution willing to know through what channel it can most effectively reduce the level of individuals' food waste. Note that such a ranking cannot be established based on a comparison of the results from existing studies as some of them are qualitative (e.g. Exodus, 2007; Graham-Rowe et al., 2014; Brook Lyndhurst, 2007) while others are quantitative (e.g. Vissenchers et al., 2016; Principato et al., 2015; Parizeau et al., 2015; Stancu et al., 2016; Stefan et al., 2013) and do not measure individuals' characteristics and behaviors in a homogenous way.

Finally, we not only look at the relationship between food risk perceptions and food waste behaviors, but also consider risk preferences in the health/safety domain and in

general. The latter variable is potentially an interesting one to include in future studies on the topic of food waste as it is easily measurable (general risk preferences can be reliably estimated by one single question (Dohmen et al., 2011)), and has been extensively studied.

The remainder of this paper is structured as follows: in section 3.2, we explain in detail our study design. In section 3.3, we describe our survey measures and provide some descriptive statistics of the data collected, while in section 3.4 we present the results of our data analyses. Section 3.5 offers a discussion of our findings and the final section concludes.

3.2 Method

The present study is based on data collected in the context of Brun and Weber (2018)'s lab-in-the-field experiment. Over the lunch break, they invited students from the University of Zurich to take part in a 20 min study on food evaluation. Their participation was rewarded with half of a packaged egg sandwich and 10 CHF. If they wanted to participate and fulfilled the eligibility criteria,[23] they were invited into a classroom where they received their sandwich half. They were asked to eat it, rate its taste, and provide their willingness-to-pay (WTP, hereafter) for receiving their sandwich's second half. While eating, they completed an extensive questionnaire that aimed at measuring their behaviors towards food waste and several personality traits. If they had not entirely eaten their first sandwich half by the end of the study, they were invited to either finish it or give it back to the experimenter (such that he could weigh it). At the end of the experiment, a die roll together with the subjects' reported WTP determined whether (and at what price, p) they would buy the second sandwich half. If the purchase took place, the subjects received their second

[23] To participate, a person should not have any allergies, intolerances or diet restrictions that would prevent them from eating the sandwich, should understand written English, and should not have previously participated in the study.

sandwich half and *10 – p* CHF, otherwise, they were simply given their full participation fee of 10 CHF.

This was a 2x2 between-subjects design where two dimensions were varied: subjects received a sandwich half that had been produced either 1 (*Fresh*) or 7 (*Old*) days before the experiment took place and they were either informed about the production date of their sandwich (*Salient*) or not (*Blind*).[24] While Brun and Weber (2018) analyze the effects of freshness information provision on participants' gustatory experience, here we focus our attention on the data collected via the questionnaire to better understand people's food waste practices. We then compare these data to subjects' *actual* tendency to waste food in the experiment, as measured by the weight of their potentially unfinished first sandwich half when leaving the experiment, and by their WTP to prevent their second half from being wasted.

3.3 Description and analysis of survey measures

Our sample consists of 258 students[25] who are between 18 and 35 years old, the median age being 22. Women comprise 52%, 79% are Swiss (potentially with an additional nationality), and 48% study natural sciences[26] (as a primary or secondary subject). In the following, we describe the data collected on reported individuals' behaviors, beliefs, and attitudes towards food waste reduction, as well as risk preferences.

3.3.1 Behaviors towards food waste

Quantity of food trashed: Participants indicated the percentage of what they buy or cook that ends up being thrown away (ca 0%, <10%, <20%, <30%, <40%, <50%, >50%,

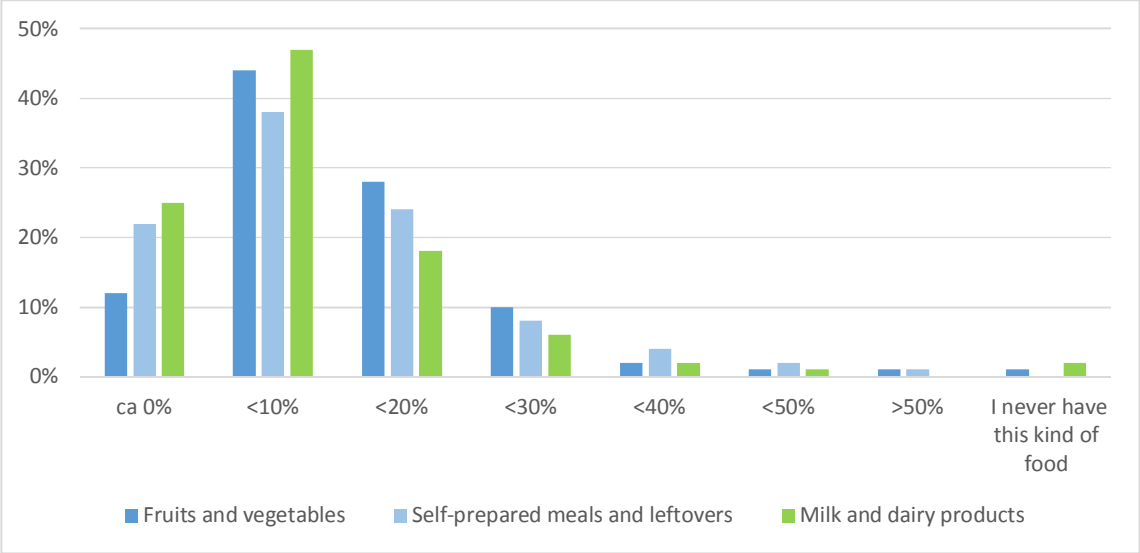
[24] For more details about the experimental design, please refer to Brun and Weber (2018).

[25] The number of observations in each condition is as follows: $N_{Blind\text{-}Fresh}$=65, $N_{Blind\text{-}Old}$=63, $N_{Salient\text{-}Fresh}$=65 and $N_{Salient\text{-}Old}$=65.

[26] Natural sciences typically include: environmental sciences, physics, biology, chemistry, health sciences, earth sciences…

or they could also indicate "I never have this kind of food") for different product categories (fresh fruits and vegetables, self-prepared meals and leftovers, and milk and dairy products). The distribution of the percentage of food trashed is very similar across the three types of food (see Figure 6), with the median participant reporting throwing away less than 10% of the food he bought or prepared (conditional on having this type of food).[27]

Figure 6: Distribution of the percentage of the (edible part of) purchased or cooked food that ends up being thrown away



Notes: N=258.

Behaviors towards old food: Subjects were provided with a list of eleven descriptions of old food items. For each food type, they indicated how likely they would be to eat it (1: *Not at all likely*, 2: *Moderately likely*, 3: *Very likely*). They were asked to assume that these items had been bought fresh, had been stored in an appropriate environment, and that they smelled and looked fine when inspecting or opening them. Table 5 (column 1) provides participants' average answers to each of the eleven questions. Participants'

[27] Using a sample of Swiss inhabitants, Visschers et al. (2016) also finds that the distribution of fruits and vegetables as well as dairy products thrown away is positively skewed, however, a much higher rate of respondents report not wasting any item of these two categories at all (34% and 51%, respectively).

reported behaviors are fairly diverse within each item, suggesting a wide heterogeneity of preferences.

Table 5: Average ratings of the likelihood and safety to eat some food items

		How likely would you be to eat this item? (1)			**How safe do you think it is to eat this item? (2)**		
nr	**Food item description:**	Not at all likely	Moderately likely	Very likely	Not at all safe	Moderately safe	Completely safe
1	A delivered (reheated) pizza with vegetables and meat that has been sitting in your fridge for 3 days	0.29	0.33	0.38	0.1	0.64	0.26
2	3-day-old leftovers of a home-made chicken-curry	0.17	0.33	0.5	0.09	0.54	0.37
3	(Reheated) leftovers of a pre-prepared veggie lasagna that you baked 4 days ago	0.24	0.35	0.41	0.11	0.58	0.31
4	An unopened plain (i.e., unflavored) yoghurt that expired 10 days ago	0.52	0.01	0.48	0.19	0.47	0.34
5	An unopened UHT milk carton that expired 2 months ago	0.51	0.29	0.21	0.38	0.41	0.22
6	An unopened pack of dry spaghetti that expired 1 year ago	0.11	0.2	0.69	0.05	0.26	0.69
7	An unopened can of tuna that expired 1 week ago	0.24	0.23	0.53	0.14	0.27	0.59
8	A raw broccoli that you bought 1 week ago	0.04	0.28	0.68	0.02	0.27	0.71
9	Hard boiled eggs that you cooked 1 week ago	0.35	0.24	0.41	0.24	0.36	0.4
10	Pre-packaged hummus opened 5 days ago	0.15	0.44	0.41	0.07	0.58	0.35
11	A jar of green pesto that you opened 2 weeks ago	0.19	0.41	0.4	0.16	0.51	0.33

Notes: $N_{nr}=258$ *for* nr *in* $\{1, 2, 4, 5, 7, 8, 11\}$ *where* nr *represents the item number. Some participants missed some questions both in columns 1 and 2, implying:* $N_3= N_{10}=257$, $N_6=256$, *and some questions in column 2, implying:* $N_9=257$.

Behaviors towards moldy cheese: Participants were further asked what they would do if they would notice some mold on a hard cheese just taken out of their fridge (1: *I eat it*, 2: *I remove only the moldy part (i.e., as little as possible) to maximize the quantity of cheese left to eat*, 3: *I cut out any part that is at all close to the mold and eat the rest*, and 4: *I throw the piece of cheese away*). Table 6 shows a wide variety of attitudes in such a situation.[28]

Table 6: Behavior towards moldy cheese

	Percentage
I eat it	0.8%
Remove only the moldy part	22.9%
Cut out if close to mold	45%
Throw it away	31.4%

Notes: N=258.

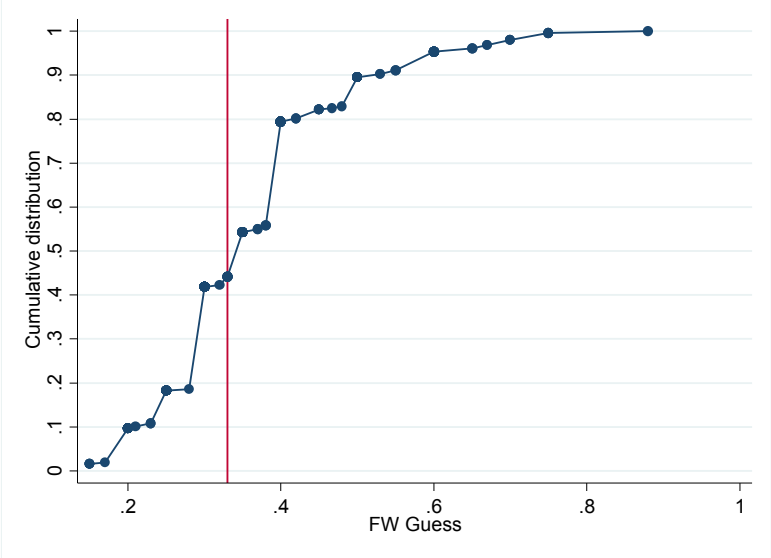
3.3.2 Descriptive norms and personal attitudes

Descriptive norms in the domain of food waste (*FW Guess*): To measure beliefs about how much food people usually throw away, subjects were asked to provide their best guess of how much of the food in households in Europe is wasted (as a percentage of the total food bought).[29] On average, participants indicated a slightly higher number than the estimation by the Food and Agriculture Organization of the United Nations of 33% (FAO, 2013; $\overline{FW\ Guess} = 0.37$, p-value of a two-sided t-test = 0.000), and the guesses were widely dispersed, as shown in Figure 7.

[28] In the case the curious reader wonders, participants with the Swiss nationality (i.e. for whom cheese is omnipresent in their culinary culture) do not display significantly different attitudes compared to foreigners (a two-sided t-test delivers a p-value=0.96).

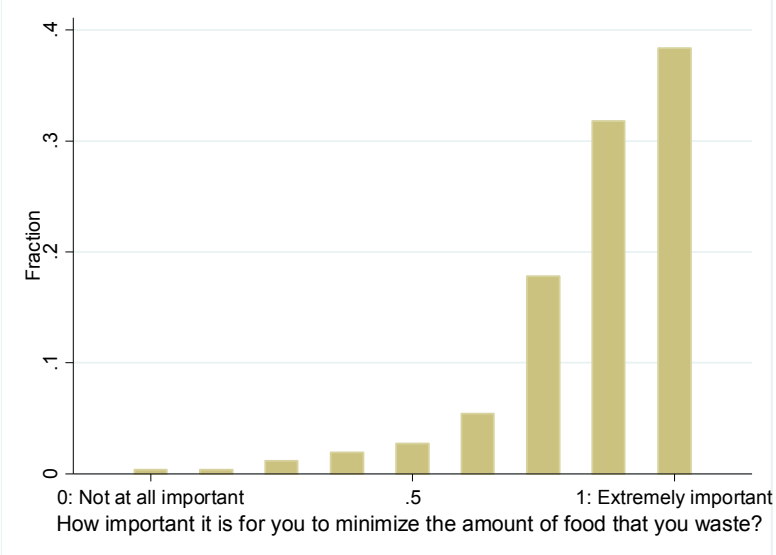
[29] Brun and Weber (2018) did not incentivize the belief elicitation to avoid making the payment phase more complex and longer. However, recent experimental evidence shows that beliefs elicited in a non-incentivized way are informative (e.g. Armantier and Treich, 2013; Trautmann and Gijs van de Kuilen, 2015).

Figure 7: Cumulative distribution of participants' guesses about the percentage of food bought by European households that is wasted (*FW Guess*)



Notes: N=258. Vertical line at 33% which is the percentage estimated by the Food and Agriculture Organization of the United Nations (FAO, 2013).

Figure 8: Distribution of participants' personal attitudes towards food waste (*Min Waste*)



Notes: N=258. Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise.

Personal attitudes towards food waste (*Min Waste*): To find out what participants think about the problem of food waste, they had to tell how important it is for them to minimize the amount of food that they usually throw away, on a scale from 1: *Not at all important* to 9: *Extremely important*. After normalizing this variable to make it lie between 0 and 1, we find that 88% of the sample reported a value equal or above 0.75 (see Figure 8 for the whole distribution). The issue of food waste thus seems to generally matter to the subjects.

3.3.3 Risk preferences

Risk preferences were evaluated in different domains using self-assessment questions. Note that all our analyses are performed on recoded measures of risk preferences such that they take values between 0 and 1.

Risk preferences in general (*Risk General*): Participants' general risk preferences were evaluated using the same question as in Falk et al. (2017), namely: *Please tell us, in general, how willing or unwilling you are to take risks?* (on a scale from 1: *Completely unwilling to take risks*, to 9: *Very willing to take risks*). A two-sided t-test indicates that participants tend on average to be slightly risk seeking ($\overline{Risk\ General} = 0.54$, p-value=0.009). However, general risk preferences vary extensively among participants as illustrated by Figure 9.

Risk preferences in the food domain (*Risk Food*): Participants indicated, for each of the same eleven old food items described in Table 5 (see section 3.3.1), how safe they think it is to eat it (1: *Not at all safe*, 2: *Moderately safe*, 3: *Completely safe*). Table 5 (column 2) shows that beliefs regarding the safety of old food varies substantially across items and individuals. We average the eleven answers within each individual to obtain an aggregate measure of risk preferences in the food domain (*Risk Food*). On average, subjects

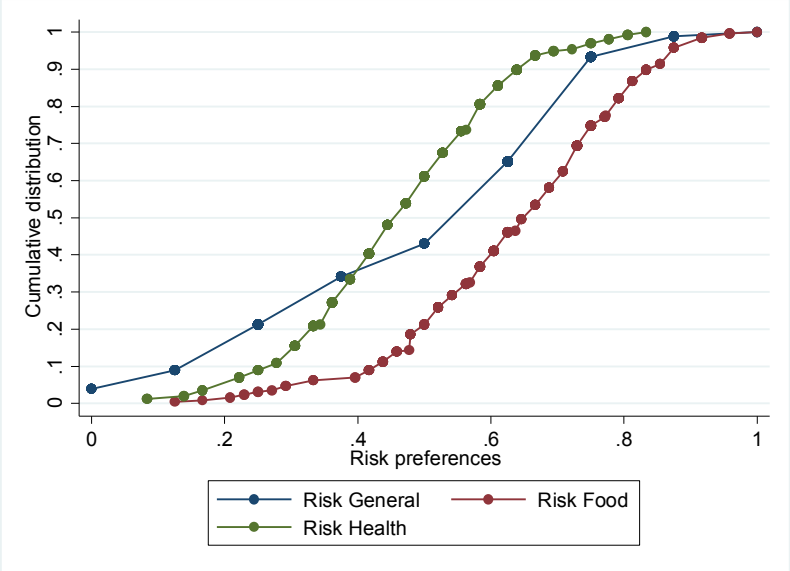
tend to believe that these food items are rather safe, with 75% of the average ratings lying above 0.5 (see Figure 9 for the cumulative distribution of *Risk Food).*

Risk preferences in the health/safety domain (*Risk Health*): An adapted version of the 9-items scale elaborated by Weber et al. (2002) is used to estimate individuals' degree of risk-taking in the health/safety domain. Respondents must indicate how likely they are to engage in each of nine different unhealthy or unsafe activities using a scale from 1: *Extremely likely*, to 5: *Extremely unlikely* (the list of items is provided in Table B2, Appendix B). We calculate the average rating provided to the nine items within individuals to obtain a measure of individuals' risk preferences in the health/safety domain (*Risk Health*). A two-sided t-test reveals that individuals in our sample are slightly risk averse when it comes to health or safety ($\overline{Risk\ Health} = 0.47$, p-value=0.001), but the distribution of risk preferences in this domain remains wide (see Figure 9).

These three measures of risk preferences positively and significantly correlate with each other (see Table 7). In particular, the correlation between *Risk General* and *Risk Food* of 0.3 (p-value <0.01) indicates that individuals' risk preferences in the food domain are not necessarily (solely) driven by a lack of expertise in evaluating the safety of old food items but also by more general preferences for risk.

Additional information on individual characteristics was collected as it may influence participants' reported and actual behaviors towards food waste. More precisely, participants were asked whether they had already experienced a serious case of food poisoning in the past (*Food Pois.*), at what frequency they prepare meals using raw ingredients (*Prep. Meals*), whether they had already eaten the type of sandwich received in the experiment (*Already Eat.*), and they also provided their level of hunger before starting the experiment (*Hungry*). Summary statistics for these variables as well as for the ones described throughout this section are available in Appendix B, Table B1.

Figure 9: Cumulative distribution of risk preferences in general, in the food domain and in the health/safety domain



Notes: N=258. Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food domain and in the health/safety domain, respectively.

Table 7: Correlation between the three measures of risk preferences

	Risk General	Risk Food	Risk Health
Risk General	1	-	-
Risk Food	0.298***	1	-
Risk Health	0.429***	0.302***	1

*Notes: N=258. Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food domain and in the health/safety domain, respectively. *p<0.1, **p<0.05, ***p<0.01.*

3.4 Results

In the following, we first construct an index – *Food Waste* – allowing us to classify individuals from small to large producers of food waste, and then investigate which of these types are more prominent across different demographic characteristics. Second, we look at whether risk preferences, descriptive norms, and subjects' personal attitudes towards food waste reduction can predict individuals' tendency to waste food. We hypothesize that participants who are not especially risk averse, who believe that people, on average, do not throw much edible food away, and/or who are concerned about the issue of food waste tend to dispose of less edible food. We also look at which of these personal characteristics best predict our construct. Finally, we test if our measure of reported food waste practices can predict subjects' *actual* behaviors in the experiment in which they participated. We expect individuals who reported wasting little food in the survey to be more likely to finish their first sandwich half in the experiment, and to be willing to pay more money to prevent their second sandwich half to be discarded.

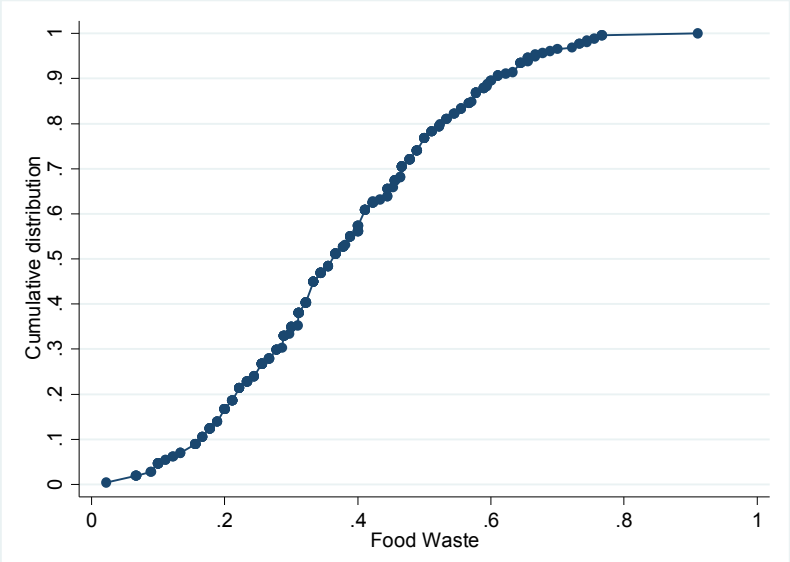
3.4.1 Construction and descriptive analysis of *Food Waste*

We create a measure of individuals' tendency to waste food by calculating a simple arithmetic mean of all 15 items[30] capturing subjects' reported behaviors towards food waste (see section 3.3.1). We obtain a new variable – *Food Waste* – that we normalize to lie between 0 and 1 such that *Food Waste* is 1 when a person wastes a lot of food, and 0 when she wastes very little food. The cumulative distribution of *Food Waste* shows a fair heterogeneity of behaviors within our sample (see Figure 10). On average, participants tend not to throw too much food away ($\overline{Food\,Waste} = 0.38, Median(Food\,Waste) = 0.37$).

[30] 4 participants missed one out of these 15 questions, and 7 other subjects answered "I never have this kind of food" when asked about the percentage of food (i.e. fruits and vegetables, self-prepared meals and leftovers, or milk and dairy products) that they throw away. In these 11 cases, *Food Waste* was obtained by computing an arithmetic mean over the 14 remaining answered questions.

In Table B3 in Appendix B, we provide the average value of *Food Waste* for different individual characteristics. Men, individuals with the Swiss nationality and students in natural sciences produce slightly less food waste than women, foreigners, and students in other study fields, respectively. Although one could expect that individuals who experienced a serious case of food poisoning in the past are more conservative when it comes to eating potentially unsafe food, we do not find evidence for this pattern. While there is no existing research on the three latter characteristics, previous work found similar results regarding the relationship between food waste behaviors and gender (Visschers et al., 2016, Secondi et al., 2015). Also, as found in previous studies (Graham-Rowe et al., 2014; Brook Lyndhurst, 2007), the subjects who report cooking often throw away on average less food than those who declare preparing meals less regularly.

Figure 10: Cumulative distribution of participants' reported behaviors towards food waste



Notes: N=258. Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food.

3.4.2 Determinants of *Food Waste*

In the following, we look at what are the main predictors of individuals' behaviors towards food waste by running regression analyses. We then compute coefficients of partial determination to find out what individual trait can best explain variations in subjects' tendency to discard edible food.

Table 8 displays the coefficient estimates from linear regressions of our measure of individuals' reported behaviors towards food waste on individuals' risk preferences in general (columns 1 and 2), in the food domain (columns 3 and 4) or in the health/safety domain (columns 5 and 6) as well as on descriptive norms and on attitudes towards food waste generation. Kruskal-Wallis H tests indicate that *Food Waste* as well as subjects' personality traits are properly randomized across the four conditions to which individuals were allocated in Brun and Weber (2018)'s experiment, except for individuals' risk preferences in the health/security domain (p-value=0.08, see Table B4 in Appendix B). As a matter of precaution, we therefore add dummy variables for treatment in each model (i.e. *Old, Salient* and *Salient*Old*). In columns 2, 4, and 6 we also control for additional individual characteristics (i.e. age, gender, people's level of hunger and whether they already experienced a serious case of food poisoning in the past).

Independently of adding all these variables, all three individual traits significantly explain subjects' reported behaviors towards food waste.[31] Individuals who take risks (in general, in the food domain, or in the health/safety domain) or who think that food waste is an important problem report wasting significantly less food ($\hat{\beta}_{Risk\ General} < 0$, $\hat{\beta}_{Risk\ Food} < 0$, $\hat{\beta}_{Risk\ Health} < 0$, p-values < 0.01; $\hat{\beta}_{Min\ Waste} < 0$, p-value$<0.05$). Descriptive norms are related to individuals' food waste practices in an intuitive direction: subjects who believe

[31] Note that the inclusion of sessions' fixed effect does neither affect the significance, nor the magnitude or the sign of the coefficients in a significant manner.

Table 8: Relationship between reported food waste behaviors and several individual traits

Dependent variable:	Food Waste					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk General	-0.198***	-0.197***				
	(-4.51)	(-4.46)				
Risk Food			-0.770***	-0.768***		
			(-20.23)	(-20.22)		
Risk Health					-0.257***	-0.259***
					(-3.97)	(-3.68)
FW Guess	0.280***	0.284***	0.160***	0.161***	0.309***	0.321***
	(3.49)	(3.48)	(2.85)	(2.79)	(3.65)	(3.73)
Min Waste	-0.137**	-0.141**	-0.150***	-0.152***	-0.168**	-0.171**
	(-2.07)	(-2.15)	(-4.08)	(-4.13)	(-2.43)	(-2.50)
Hungry		-0.00352		-0.00270		-0.00439
		(-0.62)		(-0.83)		(-0.77)
Food Pois.		0.0191		0.0112		0.0152
		(0.85)		(0.79)		(0.66)
Female		0.00278		0.00477		-0.00587
		(0.14)		(0.35)		(-0.27)
Age		0.00236		-0.000743		0.00152
		(0.76)		(-0.40)		(0.48)
Constant	0.525***	0.488***	0.944***	0.970***	0.554***	0.542***
	(7.17)	(4.57)	(20.51)	(15.69)	(6.62)	(4.65)
N	258	258	258	258	258	258
R-sq.	0.180	0.187	0.656	0.658	0.157	0.162

*Notes: Coefficient estimates of linear regression models. Dependent variable: Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female = 1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, and =0 if received a 1-day-old sandwich, and Salient=1 if was informed about the production date and 0 if did not know it. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

that people in Europe waste a lot of food throw more food away themselves ($\hat{\beta}_{FW\ Guess} > 0$, p-values<0.01).

To check for the robustness of these results, we run the same analyses with two alternative constructions of *Food Waste* – *Food Waste A* and *Food Waste B* – that aggregate the same information as *Food Waste* but in a different manner. While *Food Waste* equally weighs the 15 items capturing individuals' behaviors towards food waste, namely, the quantity of food trashed (3 items), behavior towards old food (11 items), and behavior towards moldy cheese (1 item) (see section 3.3.1), *Food Waste A* equally weighs the average behavior in each of these three sub-categories. *Food Waste B*, in contrast, consists of an average of two means: the mean of the three items measuring people's quantity of food trashed (*Food Trashed*), and the mean of the 11+1=12 remaining items capturing individuals' behaviors towards old food[32] (*Dislikes Old Food*). *Food Waste B* therefore captures people's behaviors towards food waste by giving an equal importance to their reported food discarding practices and their likelihood to eat old food.[33] Independently of the alternative construct considered, regression results are similar to the ones obtained with *Food Waste* (see Tables B7 and B8 in Appendix B).

We further compute coefficients of partial determination (CPD hereafter) for each individual trait to assess which of them can explain the largest part of variations in *Food Waste*. The estimated CPDs are reported in Table 9, where risk preferences are either captured by *Risk Food* (column 1), *Risk General* (column 2) or *Risk Health* (column 3). Independently of the domain in which risk preferences are measured, this individual

[32] We consider here moldy cheese as an old food item and therefore treat the question related to this product in the same way as the 11 ones about behaviors towards old food.

[33] *Food Waste B* is another meaningful way to capture individuals' behaviors towards food waste since the two variables that it equally weighs, *Food Trashed* and *Dislikes Old Food*, relate differently to subjects' individual traits (see regression results in Tables B5 and B6, and coefficients of partial determination in Table B9 in Appendix B).

Table 9: Coefficients of partial determination based on the regressions of *Food Waste* on individual traits

Dependent variable:	Food Waste		
	(1)	(2)	(3)
Risk Food	0.621***	-	-
Risk General	-	0.088***	-
Risk Health	-	-	0.063***
FW Guess	0.037***	0.044***	0.054***
Min Waste	0.063***	0.023**	0.034***

Notes: N=258. Coefficients of partial determination of regressing Food Waste on Risk Food (col. 1), Risk General (col. 2) or Risk Health (col. 3) as well as FW Guess and Min Waste. Dependent variable: Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, Risk Food and Risk Health are in [0.1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food- or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise. $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

characteristic significantly explains a larger part of individuals' tendency to waste food than the two other traits. While the very high CPD of *Risk Food* may reflect its similar construction to *Food Waste,*[34] *Risk General,* which is based on a single question, still reduces the sum of squared residuals (SSRs hereafter) substantially when added to the regression model ($CPD_{Risk\ General} = 0.09$, p-value<0.01). This latter variable even outperforms the marginal contribution of *Risk Health* that is based on a 9-items scale. The introduction of descriptive norms (*FW Guess*) or individuals' attitudes regarding food waste reduction (*Min Waste*) significantly decrease the SSRs when added to any of the three models, but to a smaller extent (the reduction of SSRs lies between 3.7% and 5.4% for *FW Guess*, and between 2.3% and 6.3% for *Min Waste*, depending on the model considered). When considering *Food Waste A* and *Food Waste B*, risk preferences remain the best predictor of

[34] These two indices rely on the evaluation of the same food items listed in Table 5.

variations in people's behaviors towards food waste, except when considering risk preferences in the health/security domain. In this case, *FW Guess* as well as *Min Waste* tend to outperform it (see Table B10 in Appendix B).

In brief, we find that the three individual traits that we consider are valid predictors of individuals' reported tendency to discard edible food. In particular, risk seeking subjects, those who think that a low percentage of the food purchased by consumers is trashed, and those who consider food waste as an important problem report throwing less food away. We also find that risk preference is the factor that can best explain variations in subjects' reported behaviors towards food waste.

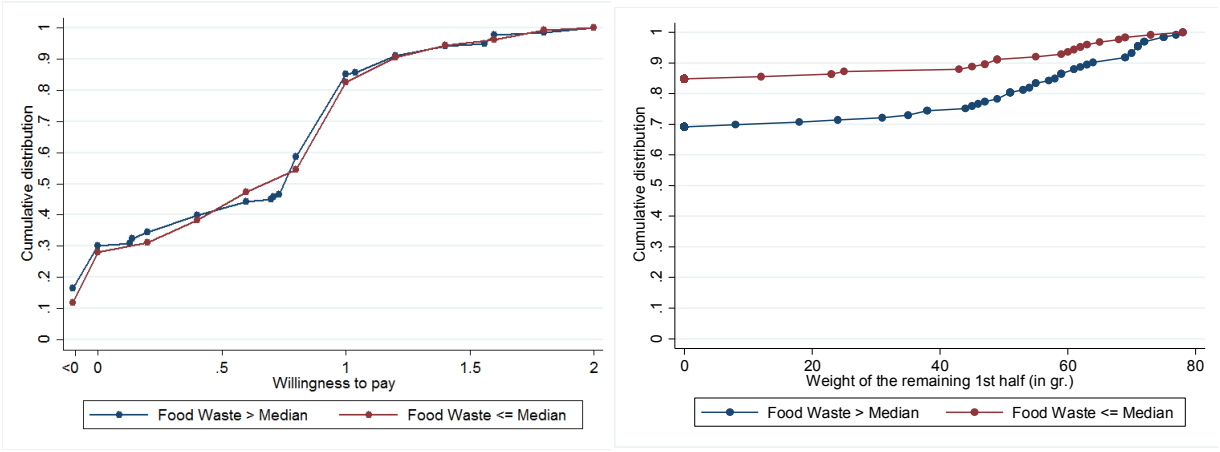
3.4.3 Validation of *Food Waste*

In this section, we look at whether individuals' reported behaviors towards food waste are informative of their *actual* behavior in a real-choice situation. As mentioned in section 3.2, participants took part in an experiment where each of them received a sandwich half that they had to eat and evaluate while answering a survey. We use their willingness-to-pay (WTP hereafter) for the second sandwich half, and the weight of their potentially unfinished first half as an approximation of *actual* behaviors towards food waste. Indeed, if a subject does not like throwing away food, she would be willing to pay a high price to prevent that her second sandwich half ends up in the trash can, and would be more likely to eat her first sandwich half entirely.

Figure 11 displays the cumulative distributions of subjects' WTP for their second sandwich half (a), and of the weight of their first half when leaving the experiment (b), with respect to whether they reported a low ($Food\ Waste \leq Median(Food\ Waste)$) or a high ($Food\ Waste > Median(Food\ Waste)$) tendency to waste food. There is no visual difference between the distributions of WTPs across the two groups. This impression is supported by a two-sided Kolmogorov-Smirnov test (p-value=0.99). When considering each

treatment group separately, there is also no clear visual difference between the two distributions (see Figure B1 in Appendix B), and the p-value of two-sided Kolmogorov-Smirnov tests is above 0.5 for each treatment group considered. However, Figure 11(b) shows that subjects reporting a low tendency to waste food ($Food\ Waste \leq Median(Food\ Waste)$) behave accordingly in the experiment by leaving behind a lower amount of unfinished sandwich than their peers.[35] This difference in distributions is (weakly) significant (the p-value of a two-sided Kolmogorov-Smirnov is 0.081) and a two-sided t-test[36] of the means indicates that the weight of sandwich leftovers is 9.1 grams higher when *Food Waste* is above its median than when it is below it (p-value = 0.003).

Figure 11: Cumulative distributions of subjects' willingness-to-pay for the second sandwich half (a) and of the weight of their unfinished first half (b) with respect to their reported food waste behaviors



(a) **(b)**

Notes: N=258. Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food.

In order to cleanly estimate the relationship between reported- and actual food usage behavior, we run regression analyses (see Table 10) to control for several individual

[35] We do not report this graph for every treatment group separately since only 60 participants did not finish their sandwich, translating into only 11 to 19 observations per condition.

[36] This test is performed on the whole sample (258 observations).

Table 10: Relationship between individuals' actual- and reported food waste behaviors

Dependent variable:	WTP weakly pos. (1/0)	WTP	WTP weakly pos. (1/0)	WTP	First half entirely eaten (0/1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Food Waste	-0.174	-0.104	0.124	-0.0457	-0.506***	-0.347***
	(-0.31)	(-0.71)	(0.22)	(-0.31)	(-3.53)	(-2.73)
Already Eat.			0.419**	0.111*		0.133***
			(2.01)	(1.78)		(3.00)
Hungry			0.151***	0.0207		0.0540***
			(3.04)	(1.38)		(4.86)
Food Pois.			-0.0115	0.0669		-0.00987
			(-0.05)	(1.12)		(-0.20)
Female			-0.427**	-0.0546		-0.209***
			(-2.09)	(-0.91)		(-4.70)
Age			-0.0448	-0.00241		-0.00724
			(-1.50)	(-0.26)		(-0.93)
Constant	1.028***	0.966***	1.119	0.824***		
	(3.48)	(11.18)	(1.48)	(3.59)		
N	258	183	258	183	258	258
Pseu. R-sq.	0.013		0.101		0.049	0.228

*Notes: Columns 1, 3, 5 and 6: Estimated marginal effects of Probit models. Columns 2 and 4: Estimated marginal effects of a truncated regression at 0. Columns 1 to 4 represent two hurdle models. Dependent variable: col. 1 and 3: WTP weakly pos.=1 if WTP>=0, and =0 if WTP<0; col. 2 and 4: WTP in [0; 2.2] CHF; col. 5 and 6: First half is entirely eaten =1 if the subject finished to eat his first sandwich half, and 0 otherwise. Independent variables: Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food; Already Eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female =1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, =0 if received a 1-day-old sandwich and Salient=1 if was informed about the production date and 0 if did not know it. McFadden Pseudo R2 are provided. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

characteristics[37] (columns 3, 4 and 6), and for the treatments to which subjects where allocated in the experiment (all columns).[38] To account for the two clusters of observations at $WTP < 0$ (37 observations) and at $WTP = 0$ (38 observations), we run a hurdle model where we regress the decision to submit a weakly positive willingness-to-pay (columns 1 and 3) and the willingness-to-pay (conditional on being weakly positive, columns 2 and 4) on *Food Waste,* using Probit and truncated regression models, respectively. Columns 5 and 6 show the estimated coefficients of Probit regressions of individuals' decision to finish or not their sandwich on our measure of individuals' behaviors towards food waste. [39]

We find that individuals' reported behaviors towards food waste are not predictive of how much they are willing to pay to prevent their second sandwich half to be discarded (the p-value of $\hat{\beta}_{Food\ Waste}$ is never below 5% in columns 1 to 4). Estimation of a Tobit model (on the restricted sample of weakly positive willingness-to-pay's, see Table B11 in Appendix B), and of a hurdle model with *Food Waste A* and *Food Waste B* as explanatory variables (see Table B12 in Appendix B) confirms this conclusion. However, *Food Waste* is predictive of subjects' decision to finish or not their sandwich: individuals who report throwing a lot of food away (*Food Waste*=1) are 35% less likely to eat their first sandwich half entirely (p-value <0.01, column 6). This result is confirmed by simple linear regression analyses (see columns 1 and 2 in Table B13, Appendix B) or when considering our alternative measures of reported food waste practices (*Food Waste A* and *Food Waste B*) using a Probit model (see columns 3 to 6 in Table B13, Appendix B).

[37] We do not only control for demographics, but also for variables that are likely to correlate with participants' WTP and with their amount of sandwiches' leftovers, namely: whether subjects had already eaten this type of sandwich in the past (Already Eat.), their current hunger level (Hungry), as well as whether they had experiences food poisoning in the past (Food Pois.).

[38] Adding sessions fixed effects do neither change the magnitude, nor the sign or the significance of coefficients substantially.

[39] Provided that only 60 out of 258 individuals did not entirely eat their sandwich, we focus our analysis on explaining the decision to finish or not the sandwich, rather than on predicting the weight of eventual leftovers.

To summarize, our measure of reported behaviors towards food waste is partly predictive of subjects' actual behavior in the experiment. Individuals who declare throwing little food away are not more willing to pay to prevent their second sandwich half to be discarded. However, they are significantly more likely to finish the sandwich half that they have started eating. While several reasons can explain the partial validation of *Food Waste*, one should mention that this index is a *general* measure of reported behaviors towards food waste. However, individuals' actual behaviors were observed in the specific environment of an experiment with decisions to make about a single food item. This context is not representative of the broad range of situations and choices that can lead to disposing of food on a typical day and may therefore explain the limited predictive power of our construct.

3.5 Discussion

The aim of the present article was to better understand what psychological factors drive individuals to discard edible food. Based on survey data, we find that risk preferences, descriptive norms, and individuals' attitudes towards the generation of food waste are relevant predictors of our construct of individuals' (reported) tendency to discard food. We further find that risk preferences can best explain variations in the latter measure, independently of the kind of risk preferences considered (i.e. risk preferences in general, in the food domain, or in the health/security domain). Taking advantage of existing behavioral data for the participants, we also test the validity of our index of reported food waste practices. We find some support for our construct to explain actual behaviors towards food waste in a real-choice situation.

Overall, our results may be of high interest for any institution or leader willing to address the issue of food waste at the consumer level in an efficient manner. In particular, our analyses suggest that any initiative or campaign against food waste that targets

individuals' perceived food risks may have a bigger impact if it influences descriptive norms or people's attitudes towards food waste reduction.

In the following, we first elaborate on the role played by descriptive norms for individuals' food waste behaviors. We then take a closer look at people's food risks perceptions by comparing them to actual food risks. Finally, we point out limitations of our study.

3.5.1 The role of descriptive norms

While we validate existing evidence on the role played by food risk perceptions (Neff et al., 2015; Evans, 2011, 2012; Graham-Rowe et al., 2014; Visschers et al., 2016) and attitudes towards food waste generation (Principato et al., 2015), we are the first study showing the additional relevant role of descriptive norms. Previous work suggested that people who feel morally obliged not to waste food also behave so (Visschers et al., 2016), and that those who think that their social group judges discarding food as a bad act have a stronger intention to reduce the amount of food that they throw away (Graham-Rowe et al., 2015; Stefan et al., 2013; Visschers et al., 2016; Stancu et al., 2016). We find that the simple belief that most of the population wastes little food is associated with lower amounts of discarded food. As 58% of our subjects overestimate the percentage of food wasted in Europe, an initiative aiming at correcting this belief may be a straightforward way to reduce the volume of edible food discarded by individuals holding these beliefs. Evidence indeed suggests that communicating information about the food waste issue via online media or the newspapers is negatively related to individuals' food waste generation (Principato et al., 2015).

3.5.2 Perceived food safety

Among the three types of risk preferences that we consider, we find that risk preferences in the food domain best predict our measure of food waste behaviors. This

validates existing evidence that food safety concerns are an important determinant of households' food disposal (Neff et al., 2015; Evans, 2011; Graham-Rowe et al., 2014; Visschers et al., 2016). In this context, the question of whether individuals hold reasonable beliefs regarding the safety of old food products is appealing. To assess the actual safety of the food items presented in Table 5 (see section 3.3.1), we asked for the opinion of a Swiss association of consumers and of a food scientist. We also looked at two reliable summary reports about food's shelf-life (FRC, 2012 and Rivet-Bonjean, 2018) and researched what the Swiss law prescribes in terms of products' bacterial content and of hygiene standards in the production of food. In Table B14 (see Appendix B), we provide the results of our investigations as well as subjects' evaluation of each food item's safety. We find evidence that for many food items, subjects' safety assessments are overly conservative. While an expired unopened plain yoghurt / UHT milk carton / pack of dry spaghettis / can of tuna, as well as a week-old raw broccoli are completely safe to eat (provided that they look and smell fine), 66%, 79%, 31%, 41% and 29% of the subjects, respectively, rated these items as *not at all safe* or *moderately safe* (see Table B14 in Appendix B). Also, when considering a hard cheese with some mold on it, more than 30% of the participants would throw the entire cheese away (see Table 6 in section 3.3.1), while removing only the moldy part would be sufficient to make the cheese safe to eat.[40] In addition to revealing people's conservative beliefs, these results also indicate that individuals do not rely enough on their senses to evaluate food,[41] and are not aware of the existence of different date labels (or of their meaning) even on common food products. For example, dried spaghettis are labelled with a "best before" date, implying that *only* their quality may deteriorate once the date is

[40] This statement represents the opinion of a food specialist working for one of the biggest Swiss supermarket and of a nutritionist dietician who published an article in a recognized French journal specialized in nutrition and health (Rivet-Bonjean, 2018).

[41] They were explicitly told to imagine that each item from Table 5 looks and smells fine when inspecting them.

passed, not their safety level. Therefore, a simple sensory evaluation would be sufficient to decide whether or not to consume them.

In our study, we do not only find that people's food risks perceptions are generally overestimated, but also that they are highly correlated with the tendency to waste food. Therefore, educating individuals to improve their ability to disentangle old- from hazardous food, by using their senses and correctly interpreting the various date labels, may be an effective way to help reduce the amount of food daily thrown away by individuals.[42]

3.5.3 Limitations

Our study has a few limitations. First of all, participants' general behaviors towards food waste are self-assessed. This is problematic for several reasons. First, individuals' reported amount of food trashed is subject to measurement errors as people may be overly influenced by the amount of food they recently discarded or they may simply not remember precisely the amount of food that they usually throw away. Kormos and Gifford (2014) indeed showed that people do not pay much attention to their daily behaviors. Second, although subjects were explicitly informed that their anonymity would be preserved throughout the study, they may still have underreported the quantity of food that they throw away to avoid feelings of shame and because they knew that their answers were not verifiable.[43] However, note that this concern may not be too prevailing as research showed that social desirability does not have a significant effect on self-reported pro-environmental behaviors (e.g. Kaiser et al., 1999; Milfont, 2009). Also, although we constructed an index of individuals' *general* tendency to waste food, it was still able to partly predict participants'

[42] Note that such measures were also suggested by several scholars (e.g. Neff et al., 2015; Graham-Rowe et al., 2014; Principato et al., 2015; Schanes et al., 2018).

[43] Individuals' reported likelihood to eat old food items may also be influenced by feelings of shame, but the direction of the bias is unclear. Individuals' social image may be tarnished both by indicating their unwillingness or their willingness to eat an old food product: they may feel embarrassed about generating food waste in the first case, and about being perceived as a person with disgusting habits in the second case.

behavior in the very specific context of an experiment. Our measure of reported behaviors should therefore be somewhat reliable.

Another caveat applies to the interpretation of our results. Firstly, we provide *correlational* and not *causal* evidence for the relationship between the three individual traits we study and food waste practices. Among others, the different measures we use should have been collected separately to prevent subjects' potential artificial consistency across their answers. Also, our study does not allow us to rule out a potential problem of reverse causality. For instance, participants who reported discarding a high amount of food may have convinced themselves that other people also do so (by indicating a high percentage of food wasted by European households) in order to reduce eventual feelings of guilt. Second, note that our findings are not necessarily extendable to the overall Swiss population as our sample consists exclusively of young and educated individuals. At the same time, we study exactly the part of the population that is among the highest contributors to food waste (e.g. Brook Lyndhurst, 2007; Visschers et al 2016; Quested et al., 2013; WRAP, 2008b; Secondi et al., 2015). Our results may be therefore particularly relevant for the elaboration of campaigns or initiatives targeting this specific segment of society.

3.6 Conclusion

The yearly global amount of food wasted has substantially increased in the last few decades and has been shown to play a non-negligible role in climate change, food security, and economic development. In fact, this issue has even been recently set as a priority at the level of the European Union (European Commission, 2011) and at The United Nations (FAO, 2011, 2013). As most of the food discarded takes place at home, the present study aimed at providing a clearer understanding of the psychological factors driving individuals to waste food. We find that risk preferences, personal attitudes towards food waste reduction, and descriptive norms significantly explain variations in people's food waste

practices and show that risk preferences are best at doing so. Considering our findings, we believe that future interventions aiming at providing food safety guidance or at informing people of the overall percentage of food wasted by consumers may be two effective ways to combat food wastage.

Chapter 4 – Psychological Opportunity Costs: The Effect of Opportunity Costs on Post-Choice Utility

Joint with Itay Sisso, Reto Odermatt and Benjamin Scheibehenne[44]

4.1 Introduction

In today's consumer societies, people are constantly making choices and need to trade off costs and benefits of different alternatives. This ranges from every day consumer decisions, such as what to buy for lunch, to more complicated and impactful life decisions, such as what career to pursue or whom to marry. Freedom to choose generally has a very positive connotation, as choosing in principle allows to better satisfy one's needs (Schwartz et al., 2006). People even seem to intrinsically value their decision rights beyond the instrumental benefits of their choices (Bartling et al., 2014). However, freedom of choice might not always lead to beneficial outcomes. In particular, we claim that opportunity costs that are, by definition, involved in every choice context (i.e., whenever we trade off different alternatives to make a decision), negatively affect consumers' well-being. We argue that in the current consumer research, this aspect has been overlooked, even though it is likely an important determinant of consumer satisfaction and reflects a systematic deviation from rational behavior.

In this project, we assess the effect of opportunity costs, that is, the forgone utility of the rejected option, on consumers' well-being. Traditional economics considers opportunity costs only to be relevant in the decision-making process; when the decision has been made, however, these opportunity costs are considered as "sunk" and people are assumed to consume the absolute level of utility of the chosen option. In contrast to this assumption,

[44]Please cite as Sisso, Itay, Brun, Fanny, Odermatt Reto and Benjamin Scheibehenne (2018) "Psychological Opportunity Costs: The Effect of Opportunity Costs on Post-Choice Utility," Working Paper.

we consider the possibility that the opportunity costs created by the forgone alternative may reduce the utility from consuming the chosen option. Importantly, we argue that this effect is present even when the forgone alternative remains less attractive than the selected one (i.e., without experiencing regret). We denote the extent to which opportunity costs matter for post-choice utility as Psychological Opportunity Costs (POCs) and hypothesize that POCs, determined by the utility of the rejected alternative, have a negative impact on post-choice utility.

Based on our theoretical considerations, we derive two main predictions: first, we hypothesize that, conditional on obtaining the preferred option, utility resulting from a non-choice situation is higher compared to a choice situation. Second, we predict that the magnitude of the effect is positively related to the forgone utility of the rejected option. The most prominent example of a related phenomenon is the feeling of regret after people have made a choice (see, e.g., Zeelenberg et al., 2000 for a review). Also, previous literature shows that the number and diversity of options might impact consumers' well-being as moderators of regret (e.g., Sagi and Friedland, 2007). However, such related theories are not able to predict our hypotheses or explain our results. First, regret theory is a *prospective* theory about choice making while our hypotheses are retrospective since they exclusively focus on the utility experienced *after* a decision has been made. Also, regret feelings can only emerge when uncertainty about the chosen and rejected options is resolved, and one realizes that the latter would have made us better off than the former. On the contrary, we argue that POCs negatively affect the utility of choice already *before* consumption has taken place, or even in the total absence of uncertainty.

We conducted a comprehensive experiment to test the extent to which the size of opportunity costs matters for post-choice satisfaction and to distinguish POCs from established theories documented mainly in the psychological literature (e.g. regret and

disappointment). In the experiment, we used hypothetical and experienced subjective well-being as a proxy measure for utility and randomized the opportunity costs people faced, i.e., the availability and desirability of inferior option(s). The experiment had a factorial design on discrete choices with five different good categories (Ice Cream, Pen, Dish, Movie and Activity). The experiment comprised three stages: in the first stage, we elicited the subjects' preferences between six options in each of the five good categories with an on-line survey. One week later, participants were then randomly treated with high (low) opportunity costs by providing them the choice between two options personally evaluated as their best and second (sixth) best option, respectively. Participants in the control group simply received their preferred option. In the third stage, we collected data on post-*consumption* happiness with the chosen ice cream and movie clip. At the end of the laboratory session, participants first watched the selected (or allocated) video and rated it. Then, they were offered the chosen/allocated ice cream before leaving and had to indicate their derived satisfaction from consuming it in an on-line survey on the same day of the experiment. Note that each subject received her allocated/chosen good only for these two categories as the consumption of the ice cream and the viewing of the video were possible at a clear point in time (i.e., on the same day of the experiment), logistically enabling the measurement of subjects' post-consumption utility. Also, we were financially constrained and these goods were the least expensive ones.

The collected data provides mixed support for our hypotheses. We find that post-choice satisfaction is significantly reduced when individuals are treated with opportunity costs: they express a lower happiness level with the chosen option compared to individuals in the control group. However, the size of this effect is independent of the attractiveness of the rejected option. Also, when considering the satisfaction *after* consumption, we do not find evidence that POCs matter.

Note that our findings for post-choice utility cannot be explained neither by regret nor by disappointment since the uncertainty about all alternatives had not been resolved. Indeed, subjects' happiness with their choice was reported *directly* after they made their decision so that they had not yet received any additional information on *any* good from the same category. At the time of the utility measurement, subjects could have therefore not felt regretful or disappointed with their choice since they still did not know if another choice would have been better or if the chosen option was as good as expected, respectively. Also, disappointment theory predicts effects that are orthogonal to our treatment since *every* subject received her favorite option. However, subjects whose outcome was determined by their choice may have experienced decision difficulty or may have *anticipated* the regret they might feel later on with their decision. Our design does not allow to confidently exclude these two channels and future research is needed to address these caveats.

The remainder of this paper is structured as follows: section 4.2 provides theoretical considerations and hypotheses, gives an overview over previous literature and shows the distinction of our concept from alternative theories. Section 4.3 describes the study design, section 4.4 displays the results, section 4.5 offers a discussion of our findings and the last section concludes.

4.2 Theoretical considerations and distinction from alternative concepts

4.2.1 Theoretical considerations and hypotheses

To better understand the concept of psychological opportunity costs and to derive testable hypotheses, we present a formal representation of the role of POCs for individual utility in the following. Consider an individual in a choice situation between the two options A and B. In the standard economic approach, rational individuals would compare the relative (expected) utility of options A and B and decide accordingly. In a situation where

A is preferred over B ($A \succ B$), A would be chosen, and B would represent the opportunity costs of choosing A. After the (irreversible) decision would be made, however, the opportunity costs from the forgone option B would not matter anymore and the absolute level of utility from good A would be consumed:

$$U_i = u_i(A) \quad (1)$$

In contrast to this view, we argue that the opportunity costs from the forgone option B leave a psychological trace and, thus, do matter for an individual's utility derived from good A. More formally, we propose a utility model that takes into account the opportunity costs induced by the unchosen alternative B in the following way:

$$U_i = u_i(A) - \alpha_i * u_i(B) \quad (2)$$

This formulation indicates that an individual's utility $u_i(A)$ from consuming the preferred option A is potentially reduced by the opportunity costs represented by the forgone utility $u_i(B)$. Psychological opportunity costs are thereby represented by the term $\alpha_i * u_i(B)$ that enters negatively in the utility function. In our framework, we restrict ourselves to choice situations between two *goods*, i.e. where $u_i(A)$ 0 and where $u_i(B)$ 0. α_i, 0 reflects the degree to which an individual *i*'s utility is affected by the opportunity costs even after the decision is made. If $\alpha_i = 0$, opportunity costs do not matter for an individual's post-choice utility, as predicted by the standard economic model. Importantly, for any values $\alpha_i > 0$, opportunity costs reduce post-choice utility.

We follow Loomes and Sugden (1982) by defining the absolute level utility $u_i(A)$ and $u_i(B)$ as the *choiceless utility* of option A and B, respectively, i.e. the experienced utility from option A or B outside any choice context. However, unlike in regret and disappointment theory (Loomes and Sugden, 1982; Bell, 1982), our model does not include expectations and does not rely on uncertainty. Our theory suggests that even when all the utilities are certain, and even when the chosen option is and remains the strictly preferred

one, the forgone option might affect individual utility of the chosen good in a negative way. One interpretation of POCs would be that the decision maker would like to choose both options but is refrained from doing so by any kind of constraint, such that sacrificing the forgone option is construed as a "painful" loss, even if it is less desirable than the chosen option.

In line with this simple model, we formulate two main hypotheses that present the predictions regarding the appearance of psychological opportunity costs:

H1: *Psychological opportunity costs negatively affect post-choice utility.*

H1 translates into the fact that POCs enter negatively in the utility function presented in equation (2). Importantly, the hypothesis predicts that one would be better off acquiring option A in a no-choice scenario than in a choice scenario between options A and B (conditional on A being the favorite option).

H2: *The magnitude of the negative effect of psychological opportunity costs is positively related to the utility of the forgone option.*

H2 shows up in our formalization of POCs: it is a function of the utility provided by the forgone option, and predicts that, in a binary choice scenario, one would be better off forgoing an option providing a low (rather than a high) utility.

To summarize, when taking into account the general effect of POCs on utility (i.e., the terms subtracted from the choiceless utility $u_i(A)$ in equation (2)), H1 states that POCs have a negative impact on the utility, while H2 predicts that the magnitude of this effect monotonically increases with $u_i(B)$.

4.2.2 Distinction from alternative theories

In this section, we review previous research that investigated how individuals' satisfaction with their choice is potentially negatively affected by the choice set. We explain how these theories and this evidence relate to our hypotheses.

The most prominent example of a similar phenomenon is the feeling of regret after people have made a choice (for the theory, see Bell, 1982, 1985, and Loomes and Sugden, 1987; for a review of the evidence, see Zeelenberg, 1999 and Zeelenberg et al., 2000). Regretful feelings arise "[…] when one is or feels responsible for the occurrence of negative events, i.e. when a different choice would had led to a better outcome" (Zeelenberg et al., 2000). One can thus be regretful because the realized outcome of a forgone option would have been better than the obtained outcome. There are two basic requirements to experience regret: uncertainty (in the level of utility of the different alternatives) and post-choice resolution of that uncertainty (Gilovich and Medvec, 1995). Regret is positively related to the difference between the forgone (B) and chosen (A) goods' utility. On the other hand, if the chosen option turns out to be better than the rejected one, as it is the case in our theoretical considerations, the latter difference would be negative, and this negative regret is defined by Loomes and Sugden's (1982, p. 808) as *rejoice*, i.e., "the extra pleasure associated with knowing that, as matters have turned out, [one] has taken the best decision". We argue that our theory is distinct from regret theory, as we predict that POCs exert a negative effect on utility *even* when the forgone alternative is less attractive than the selected one, a situation in which regret theory would predict a positive impact on the consumer's satisfaction through rejoicing. In addition, regret and rejoicing accrue once uncertainty has been resolved as in order to experience these feelings, one has to undergo

some form of surprise.[45] Note that disappointment theory also requires the decisional environment to be uncertain as feelings of disappointment only arise when the realized utility is lower than the expected one (Bell, 1985; Loomes and Sugden, 1982). In contrast, we argue that POCs can well appear in a setting either without uncertainty (i.e. in the absence of any surprise), or where uncertainty has not yet been resolved. Thus, although regret and POCs share a comparative counterfactual nature, identifying POCs as a separate construct from regret or disappointment is crucial to fully understand the impact of choice situations on consumers' well-being.

However, some other concepts reconcile with our theory and have similar predictions. For instance, studies show that an option's attractiveness decreases through joint evaluation (compared to single evaluation) (e.g., Hsee and Leclerc, 1998; Brenner et al., 1999). However, they focus on the effect of alternatives on people's evaluations at the time of decision making and do not study how the rejected options may affect post-decision satisfaction with the selected item. While making a decision is an event that has a limited duration, the time that elapses once the decision has been made extends to infinity. Since we argue that POCs matter for *post-choice* utility, their detrimental effects may stretch over a long period of time and therefore have overall possibly worse consequences for individuals' well-being than the short-lived adverse effects of joint evaluation.

Also, the basic notion of decision difficulty predicts that choice situations entail detrimental feelings that might linger after the choice, reducing one's satisfaction (Griffin and Broniarczyk, 2010; Diehl and Poynor, 2010; Goodman et al., 2013; Greifeneder and al., 2010). While we cannot entirely rule out this channel, we can empirically test whether our treatment effect is driven by individuals who have difficulty in making decisions.

[45] Indeed, in decision affect theory (Mellers et al., 1997), which is a more recent model of regret and disappointment, the elements representing the effects of regret and disappointment are multiplied by a "surprisingness" factor, which takes the value of zero when no uncertainty applies.

Another example of a phenomenon which would have similar predictions to our second hypothesis is the long-established contrast effect (Brown, 1953), which posits that a bad option in the choice set might make the chosen option look much better, while a very good option in the choice set might have the reversed effect. However, this theory produces a different prediction from our hypothesis 1. While the contrast effect implies that participants rejecting an unattractive option would be better off than if they had not been offered an alternative at all, our theory stipulates an opposite effect: the opportunity costs produced by the rejection of the alternative option hinders utility with the chosen option.

Some additional phenomena predict effects that contradict ours (i.e., making the chosen option more attractive in a choice situation compared to a situation without a choice), such as dissonance reduction (e.g., Brehm, 1956 or Festinger, 1957), single option aversion (Mochon, 2013), the intrinsic decision value (e.g., Bartling et al., 2014), or simply the pleasure of choosing (Botti, 2004).

Finally, another line of effects called *context effects*, might include both supporting and contradicting predictions. For example, the asymmetric dominance effect (also known as the attraction or decoy effect; Huber and al., 1982) suggests that the existence of a slightly less attractive version of an option (usually referred to as the *decoy*) might increase the attractiveness of a target option. This contradicts our hypothesis that an alternative option can only (potentially) have detrimental effects for the utility of the target option through the psychological opportunity costs that it generates. However, an opposite context effect is the similarity effect (Tversky, 1972), which posits that the presence of a similar option to the chosen one might reduce the attractiveness of the latter. While this phenomenon conforms to our prediction, we do not impose any constraints regarding the similarity of the options. While we assert that rejecting an option with a similar attractiveness level as the chosen one will cause the highest level of POC, these options

might very well be quite different from each other (e.g., one may find both a fun and flashy sports car and a comfortable and practical one attractive although they are characterized by very different attributes).

4.3 Study design

4.3.1 Design overview

Our study consists of three steps: first, participants take part in an on-line survey, in which we collect information on participants' preferences among six options in each of five good categories. We then use individuals' preference orderings to randomly assign three different levels of opportunity costs in the laboratory session, taking place roughly one week after the on-line survey. We do so by letting participants select an item from a choice set consisting of their favorite option and another less appealing one whose attractiveness is exogenously varied. Varying the extent to which a person likes the latter option allows us to test the extent to which opportunity costs affect individuals' post-decision utility. After the experimental session, subjects complete a 3 min long on-line survey on the same day. This third stage allows us to measure post-*consumption* utility. Following a growing literature in happiness economics (for reviews, see, e.g., Kahneman et al. 1999, Frey and Stutzer 2002, Kahneman and Krueger 2006, or recently Odermatt and Stutzer 2018), we use reported subjective well-being as an empirical approximation of utility. In the next paragraphs, we describe the three steps of our study and the different treatments in details.

4.3.2 Stage 1: on-line survey

The on-line survey, based on the Qualtrics software (Qualtrics, Provo, UT), had to be completed at least seven days before the laboratory session took place.[46] The purpose of the

[46] We chose this seven-day gap between the survey and the laboratory session in order to reduce the salience of the different goods. We abstained from using a longer gap to avoid that people's preferences change too much between these two stages.

survey was to elicit subjects' preferences for goods within different good categories. More specifically, participants were shown, one at a time, six pictures of different goods belonging to the same good category (e.g., six ice creams of different flavors). We asked the following question to reveal people's valuation of the respective good: "Imagine that you would receive the following [good]. How happy would [this good][47] make you?". The scale ranged from 1 "Not at all happy" to 100 "Very happy".[48] The happiness ratings were then used to individually rank the six products in every good category from the preferred (first best) to the least preferred one (sixth best). This was done to individualize the treatments in the second stage, as explained in section 4.3.3. After having evaluated the six goods from each of five different good categories, participants had to rank the six goods from the respective good category from the most to the least preferred one. This allowed us to identify the most preferred good in cases same happiness ratings were attributed to two different goods within a category.[49] All these ratings were incentivized as we informed subjects that they would receive at least one good as a thank you gift at the end of the study and that evaluating each good carefully would increase the probability that they receive a good that they like.

In the survey, we also asked whether they could not eat one or more of the six dishes that they were presented because of food restrictions, we elicited several demographic characteristics (age, gender, nationalities, and field(s) of study) and we measured people's difficulty in making decisions using a shorter version of the Decision Difficulty scale by Turner et al. (2012) (see Table C3 in Appendix C for a description of this measure).

Good categories: Participants were presented six distinct goods from the following five good categories (see Tables C1 and C2 in Appendix C for a detailed description of all

[47] The term in bracket was replaced by the type of good considered.

[48] See Figure C1 in Appendix C for an overview of how this question looked like in the on-line survey for the Ice Cream category.

[49] We could have solely used subjects' rankings instead of their happiness ratings, however, the latter also serve as a measures of each option's choiceless utility, which we later use in our analyses.

thirty items, and for the 30 pictures presented to the subjects, respectively): Moevenpick ice creams of six different flavors, Caran d'Ache ball pens of six different colors, vouchers for six different meals in restaurants in Geneva, short educative and animated videos (of approximately 3 minutes length) on six different topics, and vouchers for six different activities in Switzerland. In all the good categories we used, the options differ from one another on one attribute (such as the color of the pen) or several ones (such as the many ingredients and textures of the dishes).

4.3.3 Stage 2: laboratory experiment

In the laboratory experiment, programmed with zTree (Fischbacher, 2007), the participants were randomly allocated to two different main conditions (*No-Choice* and *Choice*). Subjects in the *No-Choice* group were not exposed to any opportunity costs, while those in the *Choice* condition were treated with low and high opportunity costs (*Low POCs* and *High POCS,* respectively). In the following, we describe in detail how the design differed between the different treatments.

***No-Choice* condition**: For all five good categories separately, we showed to the participants in the *No-Choice* condition the good that they had indicated in the on-line survey to be the most preferred one in the respective good category. We informed them that they would receive (or would have the chance to receive) the good under consideration and we subsequently measured their expected happiness with the good by asking "How happy [would receiving this good] make you?", [50] measured on a scale from 1: "Not happy at all" to 100: "Very happy". The purpose of this condition was to measure individuals' choiceless utility approximated by expected happiness with their favorite option in an

[50] The terms in brackets were adapted to each goods' category, in particular, the future tense was used when talking about the ice creams and the movies, as subjects were informed that they would receive the ice cream and would watch the video at the end of the session (see Figure C2 in Appendix C for an overview of the Ztree interface for individuals in the No-Choice and in the Choice condition for the Ice Cream category).

environment in which they did not face any direct opportunity costs. Note that the order in which the good categories appeared was randomly determined.

***Choice* condition:** While subjects in the No-Choice condition did not have to make any active decisions, the participants in the Choice condition had to make a choice for every good category between two options and were thus treated with opportunity costs. In particular, participants were given the choice between the good that they indicated in the on-line survey to be the best option (first best) and another option which attractiveness was exogenously varied. We had two different treatments in the Choice condition that aimed at generating high and low levels of psychological opportunity costs (the *High POCs* and *Low POCs* treatments, respectively). In the High POCs treatment, the other option in the choice set was the one that the participant had evaluated as being the second best in the on-line survey, while in the Low POCs treatment, this option was the one that the participant liked the least out of the six possibilities (i.e., the sixth best). Each participant in the Choice condition made two (or three) decisions in the High POCs treatment, and the three (or two) remaining ones in the Low POCs treatment.[51] After each choice, we measured subjects' expected happiness with their good in the same manner as for the No-Choice group. Once all choices were made, we gave the participants a chance to modify them to make sure that they did not regret their decisions. The decisions were incentivized: for the good categories Movie and Ice Cream, the participants were able to consume the chosen good at the end of the experiment, while for the other goods, a lottery determined three participants who received their selected pen, voucher for a meal, and voucher for an activity, respectively.

At the end of the laboratory session, the individually chosen or assigned video was played on the participants' screen. After watching the video, the participants had to state

[51] The number of decisions to make in each treatment was balanced across all individuals in the *Choice* condition.

their happiness level with viewing it and rated it. Finally, each participant received the ice cream that she had selected or was allocated.

4.3.4 Stage 3: post-experiment on-line survey

After the laboratory session, subjects had to complete a 3 min long on-line survey on the same day of the experiment. In this third stage, we measured post-consumption happiness with the ice cream. We asked participants "How happy did the ice cream make you?" conditional on having at least tried the ice cream. Again, they were able to reply on a scale from 1 ("Not at all happy"), to 100 ("Very happy"). Lastly, we asked if they would have preferred another flavor to find out whether they regretted their choice.

4.4 Results

Our sample consists of students who had to participate in our study as a requirement to pass a bachelor course on consumer behavior at the University of Geneva. We conducted 25 sessions, with 1 to 6 students per session. A session lasted on average 15 minutes. Overall, out of the 132 students who completed the on-line survey, 126 of them participated in the laboratory session. Our final main dataset therefore consists of 630 observations from 126 subjects (5 observations each). Thereof, 35 students were randomly allocated to the No-Choice condition (i.e., the control group) and 91 to the Choice conditions (i.e., the High POCs and Low POCs treatments).[52] 17 subjects (all in the Choice conditions) reported that they had already watched at least one of the two videos from their choice set, 4 individuals indicated in the on-line survey that they cannot eat ice creams and 1 individual indicated that he accidentally ranked the ice cream flavors in the reverse order. These 22 observed

[52] The High POCs treatment consists of 42 (Movie), 46 (Ice Cream, Activity), and 47 (Pen, Dish) observations for the respective good category, and the Low POCs treatment consists of 44 (Pen, Dish), 45 (Ice Cream, Activity) and 49 (Movie) observations for the respective good category, summing up to 91 observations per category.

choices were therefore qualified as *non-eligible* and were excluded from the main analysis.[53] The final dataset consists of subjects between 19 and 27 years old, with an average age of 21.5, and 48% are women. Summary statistics of happiness ratings in stage 1 (*Stage 1 Happ.*), right after the allocation or after the choice was made (*Post-Choice Happ.*) and after consumption took place (*Post-Consumption Happ.*) are available in Table C4, in Appendix C.

To check that our treatments induced the desired difference in desirability between the two options (i.e., the size of opportunity costs) in the Choice conditions, we looked at the stated happiness (from stage 1) of the second or sixth best option presented to the subjects in stage 2. While the average happiness ratings of the less attractive option in the Low POCs treatment was only of 28.5 (SD=24.8) points, it was on average of 69.7 (SD=21.9) points in the High POCs treatment (which is significantly different from the Low POCs treatment, $t(453) = 18.8$, $p<.001$).[54] We therefore successfully generated decisional environments with either high or low (psychological) opportunity costs.

4.4.1 Post-choice analysis

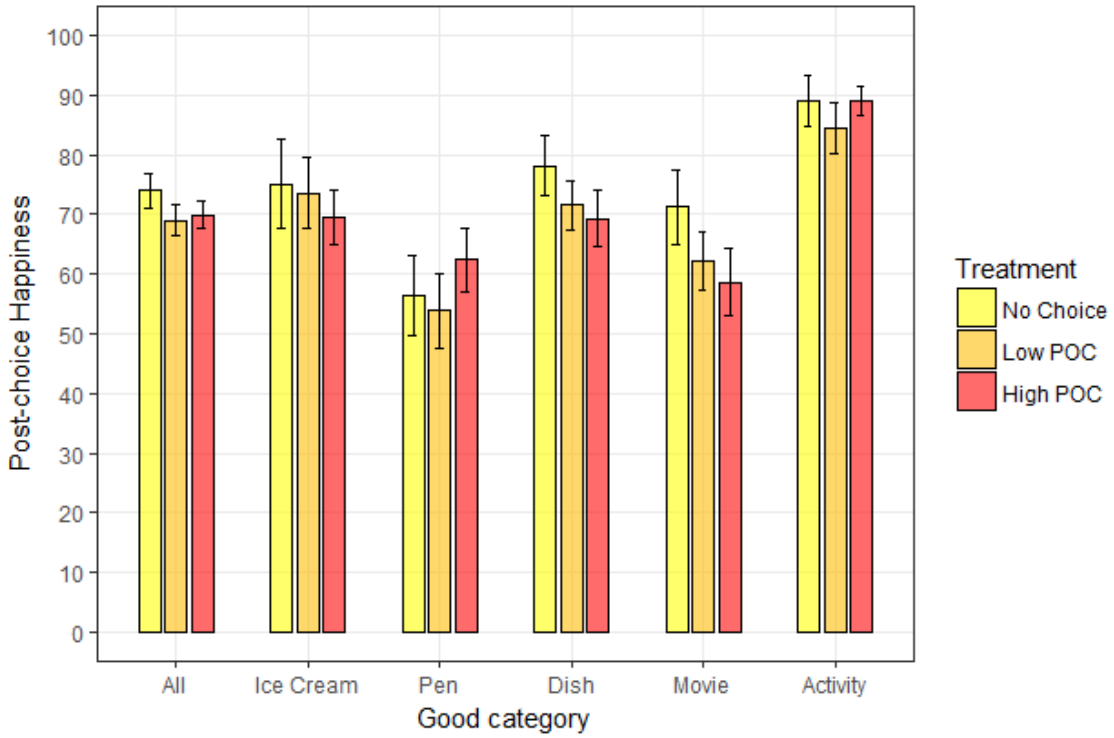
Descriptive statistics: Figure 12 shows descriptively the mean-levels of reported post-choice happiness for each good category by condition, as well as the means aggregated over all good categories. According to hypotheses H1 and H2, we expect the average happiness ratings to be higher in the No-Choice than in the Choice condition, and to be the lowest in the High POCs group. While the results for Ice Cream, Dish and Movie corroborate our hypotheses, our findings for Pen and Activity do not. When merging all good categories together, reported happiness in the No-Choice condition is higher than in

[53] 38 participants indicated that they cannot eat between 1 to 3 dishes out of the 6 presented and a participant, who indicated that he cannot eat 5 out of the 6 dishes, was randomly assigned to the No-Choice condition. These observations were all kept in our final dataset.

[54] The overall average happiness level with the first best option reported in stage 1 is of 78.0 (SD=21.9) points.

the Choice treatments ($\overline{Post-Choice\ Happ.}_{No-Choice} = 74$, $\overline{Post-Choice\ Happ.}_{Choice} = 69.5$), however, the difference is small and only weakly significant (the p-value from a two-sided t-test from a regression with standard errors clustered at the individual level is 0.067).

Figure 12: Average post-choice happiness levels for each good category separately and aggregated together



Notes: For High POCs: N_{All}=228, $N_{IceCream} = N_{Activity} = 46$, $N_{Pen} = N_{Dish} = 47$, $N_{Movie} = 42$. For Low POCs: N_{All}=227, $N_{IceCream} = N_{Activity} = 45$, $N_{Pen} = N_{Dish} = 44$, $N_{Movie} = 49$. For No-Choice: N_{All}=175, $N_{AnyCategory} = 35$. Mean of reported happiness level with the allocated/chosen option (before consumption) in the No-Choice (yellow), Low POCs (orange), and High POCs (red) condition, respectively, for each good category separately and for all good categories together. Post-Choice Happiness is in {1, 2, 3, … 100} where 1 means "Not at all happy" and 100 means "Very happy". The error bars represent the 90% CI around the mean.

We do not find a statistical difference between the average happiness ratings from the Low and the High POCs conditions ($\overline{Post-Choice\ Happ.}_{Low\ POCs} = 69.93$, $\overline{Post-Choice\ Happ.}_{High\ POCs} = 70.32$, p-value of a two-sided t-test=0.85). While Figure 12 provides descriptive insights regarding the treatment effects, we provide next the results from regression analyses to assess in a more precise way the effects of our treatments. First,

we look at whether subjects' post-choice happiness is negatively affected by the opportunity costs induced by the simple rejection of an option. In a second step, we investigate whether the size of these psychological opportunity costs is positively related to the attractiveness level of the rejected alternative.

Difference between the No-Choice and the Choice conditions (H1):

To test hypothesis H1, we analyze the difference between the No-Choice condition where subjects were *allocated* their first best option, and the Choice condition, where subjects had to choose between their first best and either their second best or their sixth best option in a multiple regression framework. The results of linear regression models are displayed in Table 11. The dependent variable is the participants' stated happiness with the chosen/allocated good in stage 2 (*Post-Choice Happ.*) and the explanatory variable is a dummy variable (*dummy*Choice) indicating whether the participant was in any of the Choice conditions or not. Importantly, we control for the happiness ratings of the chosen/allocated good from each category reported in stage 1 (*Stage 1 Happ.*). This accounts for two key features. First, it controls for possible individual-level differences in the general level of happiness with the good that is independent of the treatment. Second, for individuals in the Choice condition, the measured post-choice happiness in stage 2 consists of *both* the choiceless utility and the effect of POCs, while the happiness of the preferred good measured in stage 1 solely captures the choiceless utility. Controlling for it thus serves to estimate the effects of POCs in a more precise and reliable way. Columns 1 and 2 display the results when the observations are aggregated across all good categories together, without and with good categories' fixed effects, respectively, while columns 3 to 7 show the coefficients estimates when considering the data for each good category separately. In the first two columns, standard errors are clustered at the individual level to account for the potentially correlated happiness ratings within participants.

When considering each good category separately (columns 3 to 7), the opportunity costs generated in the Choice condition have a negative effect on participants' satisfaction with the chosen ice cream, activity, dish and movie. However, this effect is only significantly different from zero for the last two categories (for Dish: $\hat{\beta}_{dummyChoice}$=-6.62, p<0.05 and for Movie: $\hat{\beta}_{dummyChoice}$=-8.33, p<0.05) and marginally significant for Ice Cream ($\hat{\beta}_{dummyChoice}$=-5.16, p<0.1). While the average effect of the Choice condition for Pen is positive, it is small and not statistically significant ($\hat{\beta}_{dummyChoice}$=1.64, p>0.1). When considering the data pooled over all good categories, the resulting average effect of (low and high) POCs on happiness is -4.22 points (p<0.05, column 2).[55] Re-including the 22 observations that we classified as non-eligible does not affect this result significantly (see columns 1 and 2 from Table C5, Appendix C). A Tobit model accounting for the 65 cases where the dependent variable takes the maximum value of 100, produces an even more negative estimated treatment effect (see columns 3 and 4 in Table C5, Appendix C).

These results cannot be accounted for neither by regret nor by disappointment for two reasons. First, the subjects did not collect any additional information regarding the chosen/allocated or forgone options when asked to indicate their happiness with their choice (pre-consumption), which would be required for feelings of regret or disappointment to arise. Furthermore, we made sure that participants do not regret their decision: after they had selected an option for each good category, they were given a chance to modify their choices (in which case they had to indicate again their happiness with the newly chosen item).[56]

[55] Note that controlling for age and gender does neither influence the significance nor the magnitude of these findings significantly.

[56] 4 individuals changed their choice (1 in the Low POCs and 3 in the High POCs condition, respectively).

Table 11: Relationship between post-choice happiness and the treatment condition *Choice* versus *No-Choice*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:				**Post-Choice Happ.**			
Good category:	All	All	Ice Cream	Pen	Dish	Movie	Activity
***dummy*Choice**	-3.916**	-4.224**	-5.163*	1.639	-6.620**	-8.332**	-2.663
	(1.614)	(1.648)	(2.994)	(3.900)	(2.691)	(3.321)	(2.467)
Stage 1 Happ.	0.651***	0.551***	0.607***	0.452***	0.529***	0.647***	0.678***
	(0.0366)	(0.0419)	(0.0864)	(0.0730)	(0.0859)	(0.0879)	(0.136)
***dummy*Pen**		-6.734***					
		(2.154)					
***dummy*Dish**		-0.220					
		(1.561)					
***dummy*Movie**		-6.565***					
		(2.075)					
***dummy*Activity**		7.479***					
		(1.516)					
Constant	23.13***	32.24***	28.32***	27.63***	35.55***	21.34***	26.95***
	(3.33)	(4.051)	(8.177)	(5.981)	(7.721)	(7.920)	(13.479)
Observations	608	608	121	126	126	109	126
R^2	0.403	0.451	0.380	0.249	0.272	0.422	0.254
SE	Clustered	Clustered	Robust	Robust	Robust	Robust	Robust

*Notes: Coefficient estimates of linear regression models. Dependent variable: Post-Choice Happ. is the happiness level with the chosen/allocated option in the laboratory (pre-consumption) and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyChoice=1 if is in the Choice condition, and 0 if is in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyPen, dummyDish, dummyMovie and dummyActivity are the good categories fixed effects and are in {0,1} (e.g. dummyPen=1 if the post-choice happiness is reported for the Pen good category, and 0 otherwise). The Ice Cream good category represents the reference category. Standard errors in columns 1 and 2 are clustered at the individual level. Standard errors in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table 12: Relationship between post-choice happiness and the choice conditions *Low POCs* versus *High POCs*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:				**Post-Choice Happ.**			
Good category:	All	All	Ice Cream	Pen	Dish	Movie	Activity
***dummy*NoChoice**	4.983***	5.059***	4.329	3.274	7.020**	5.157	4.859
	(1.828)	(1.873)	(3.613)	(4.415)	(3.036)	(3.621)	(3.047)
***dummy*HighPOCs**	2.124	1.669	-1.607	9.507**	0.788	-7.038*	4.346
	(1.592)	(1.541)	(3.701)	(4.332)	(3.568)	(3.869)	(2.640)
Stage 1 Happ.	0.653***	0.554***	0.601***	0.457***	0.532***	0.657***	0.676***
	(0.0363)	(0.0418)	(0.0882)	(0.0704)	(0.0876)	(0.0841)	(0.136)
***dummy*Pen**		-6.686***					
		(2.147)					
***dummy*Dish**		-0.205					
		(1.576)					
***dummy*Movie**		-6.460***					
		(2.080)					
***dummy*Activity**		7.479***					
		(1.524)					
Constant	17.97***	26.98***	24.54***	24.03***	28.28***	15.34**	22.28*
	(3.299)	(3.921)	(7.967)	(5.316)	(7.317)	(6.561)	(13.18)
Observations	608	608	121	126	126	109	126
R^2	0.405	0.452	0.381	0.278	0.272	0.441	0.271
SE	Clustered	Clustered	Robust	Robust	Robust	Robust	Robust

*Notes: Coefficient estimates of linear regression models. Dependent variable: Post-Choice Happ. is the happiness level with the chosen/allocated option in the laboratory and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyNoChoice=1 if is in the No-Choice condition, and 0 if is in the Choice condition; dummyHighPOCs=1 if is in the High POCs condition, and 0 if is in the Low POCs or in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyPen, dummyDish, dummyMovie and dummyActivity are the good categories fixed effects and are in {0,1} (e.g. dummyPen=1 if the post-choice happiness is reported for the Pen good category, and 0 otherwise). The Ice Cream good category represents the reference category. Standard errors in columns 1 and 2 are clustered at the individual level. Standard errors in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

In summary, we find that individuals' satisfaction with their favorite option is lower following a choice situation – i.e., when they had to actively forgo an alternative, than following a situation without choice – i.e., when they were allocated their preferred good. Opportunity costs therefore seem to matter for utility even after a decision is made.

Difference between low- and high levels of opportunity costs (H2):

In the following, we test hypothesis H2 predicting that the magnitude of the POCs is positively related to the utility that the forgone option would provide. We test this hypothesis by looking specifically at the difference between the Low- and the High POCs conditions.[57]

We use exactly the same regression specifications as in Table 11 but including two separate dummy variables for the No-Choice and the High POCs conditions (*dummy*NoChoice and *dummy*HighPOCs, respectively) instead of the single dummy for the Choice condition (see Table 12). We do not find clear evidence for the more negative effects of high compared to low opportunity costs. When analyzing every good category separately (columns 3 to 7), the coefficient estimates for *dummy*HighPOCs are negative for Ice Cream and Movie (and marginally significant for the latter $\hat{\beta}_{dummyHighPOCs}$=-7.04, p<0.1), while the coefficients are positive for the three remaining good categories (and significant for pens, $\hat{\beta}_{dummyHighPOCs}$=9.51, p<0.05). This lack of clear support for hypothesis H2 is confirmed when considering the pooled data over all good categories (columns 1 and 2). We find that when the forgone option is the second best one, individuals do not report a significantly lower happiness level with their choice ($\hat{\beta}_{dummyHighPOCs}$=1.67, p>0.1, column 2) than when

[57] Preferences were not completely stable over time – Out of the 455 choices made in stage 2, 89% chose their best ranked option from stage 1, with 96.5% of the subjects in the Low POCs condition, and 81.6% of the individuals in the High POCs condition (a two-sided χ^2test shows that this percentage is significantly higher in the Low than in the High POCs condition, $\chi^2(1)$=25.6, p<.001). Subjects' preferences were therefore less stable when choosing between two attractive options than between an attractive and a less attractive good.

rejecting the sixth best option. This finding is confirmed by linear regressions including observations that we qualified as non-eligible and Tobit models (see columns 1 to 4 in Table C6, Appendix C). Therefore, individuals in the High POCs treatment are not significantly worse off than those in the Low POCs group as we had hypothesized. In other words, our findings suggest that the *size* of opportunity costs do not matter for post-choice utility: rejecting an attractive or an unattractive option are both equally detrimental for one's utility.

Overall, we find evidence that decision makers' post-choice utility is on average higher when they are allocated their favorite option (i.e., in an environment without opportunity costs), compared to when they have to select it from a set of two items (i.e., in an environment with high or low opportunity costs). However, the size of the opportunity costs does not matter for individuals' utility: participants who gave up an attractive item were not significantly less happy with their choice than those who rejected a less desirable one. In other words, it seems that POCs are significantly detrimental for post-choice utility, but that their size does not further amplify this negative effect.

4.4.2 Post-consumption analysis

The analysis so far focused on post-choice utility measured before receiving the goods. As mentioned in the description of the study, for two goods categories – Ice Cream and Movie – all participants were given the opportunity to consume the good that they had chosen or were allocated, and then had to report their happiness with it after having consumed it. In this section we first describe this additional data collected and then look at whether POCs matter for post-consumption utility by testing for our two hypotheses.

21 subjects reported that they had not yet consumed the ice cream by the time of the follow-up survey, and 10 did not complete the survey at all. These observations were therefore excluded from the main analysis. As in the post-choice analysis, 19 additional

observations were also removed because they were qualified as *non-eligible* (i.e., 2 individuals could not eat ice creams, and 17 had already watched at least one of the two movie clips).

Unlike the post-choice measure of happiness, post-consumption utility can be affected by regretful feelings since some uncertainty has been resolved. This is a concern as they could lead to similar findings as predicted by our theory. To address this issue, we asked participants whether they would have chosen differently after they had watched the video and consumed the ice cream, respectively. This was the case for 6 observations (7.7%) in the Low POCs treatment (3 in each the Ice Cream and the Movie categories) and for 14 observations (21.9%) in the High POCs treatment (7 in each the Ice Cream and the Movie categories). In those 20 cases, individuals believed that they would have been better off consuming the option that they rejected instead of the one that they selected. They may have therefore felt regretful about their choice and enjoyed it less. To rule out regret as an alternative explanation to our results, we thus excluded those 20 observations from our main analysis as well. As a result, 182 observations remained in the final dataset (83 observations for the Ice Cream category and 99 for the Movie category).

To test our two hypotheses, we fit the same linear regression models used for the post-choice analysis. However, this time our dependent variable is post-*consumption* happiness and we are considering only the two good categories for which we collected this measure (Ice Cream and Movie). We systematically perform three robustness checks: we run a Tobit model accounting for the 17 censored responses with a post-consumption happiness of 100, and we run linear regressions after re-including the observations that were either excluded due to *non-eligibility* or due to expressions of regret.

In the following, we first look at whether subjects' experienced utility of consuming the ice cream and of watching the short movie clip was significantly affected by the fact of forgoing an alternative (Choice condition) or not (No-Choice condition). Table 13 provides

results of regressing post-consumption happiness ratings on the treatment dummy, *dummy*Choice, for the Ice Cream (column 3) and the Movie (column 4) categories separately, and together without and with good categories fixed effects (columns 1 and 2, respectively). The regression results indicate that POCs do not affect subjects' utility after consumption occurred, independently of considering the aggregated sample or both good categories separately ($\hat{\beta}_{dummyChoice}$ is never significant in columns 1 to 4). This finding is supported by our three robustness checks (see Table C7 in Appendix C).[58] Thus, we do not find evidence that POCs significantly affect post-consumption utility after forgoing *any* option (i.e., attractive or unattractive).

Table 13: Relationship between post-consumption happiness and the treatment condition *Choice* versus *No-Choice*

	(1)	(2)	(3)	(4)
Dependent variable:		**Post-Consumption Happ.**		
Good category:	All	All	Ice Cream	Movie
***dummy*Choice**	1.340	-0.203	0.223	-0.536
	(3.532)	(3.548)	(3.786)	(4.943)
Stage 1 Happ.	0.549***	0.436***	0.440***	0.433***
	(0.0884)	(0.0897)	(0.0939)	(0.136)
***dummy*Movie**		-18.49***		
		(2.464)		
Constant	22.83***	42.77***	42.16***	24.73**
	(7.962)	(8.380)	(9.215)	(11.55)
Observations	182	182	83	99
R^2	0.207	0.339	0.233	0.125
SE	Clustered	Clustered	Robust	Robust

*Notes: Coefficient estimates of linear regression models. Dependent variable: Post-Consumption Happ. is the happiness level with the chosen/allocated option after consumption occurred and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyChoice=1 if is in the Choice condition, and 0 if is in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyMovie=1 if (post-consumption) happiness is reported for the Movie good category, and 0 if is reported for the Ice Cream good category. Standard errors in columns 1 and 2 are clustered at the individual level. Standard errors in parentheses, * p < 0.1; ** p < 0.05; *** p < 0.01.*

[58] We also obtain similar effects when controlling for gender and age.

Next, we investigate more specifically whether the rejection of an attractive option is more detrimental for post-consumption utility than forgoing an unattractive alternative. Table 14 displays the results from regressing post-consumption happiness on the treatments' dummies for the No-Choice and High POCs groups. On average, we find no significant effect of our treatment, independently of pooling or not the data across the two good categories ($\hat{\beta}_{dummyHighPOC}$ has a p-value>0.1, see columns 1 to 4),[59] and of using different model specifications or samples (see Table C8 in Appendix C).

Table 14: Relationship between post-consumption happiness and the choice conditions *Low POCs* versus *High POCs*

	(1)	(2)	(3)	(4)
Dependent variable:		**Post-Consumption Happ.**		
Good category:	All	All	Ice Cream	Movie
***dummy*NoChoice**	-1.208	0.0783	-2.048	1.950
	(3.840)	(3.752)	(4.020)	(5.675)
***dummy*HighPOC**	0.328	-0.311	-4.666	3.433
	(3.787)	(3.211)	(4.511)	(6.015)
Stage 1 Happ.	0.550***	0.436***	0.414***	0.429***
	(0.0888)	(0.0903)	(0.0994)	(0.137)
***dummy*Movie**		-18.50***		
		(2.461)		
Constant	23.99***	42.75***	46.43***	23.07**
	(7.025)	(7.183)	(8.340)	(10.36)
Observations	182	182	83	99
R^2	0.207	0.339	0.244	0.128
SE	Clustered	Clustered	Robust	Robust

*Notes: Coefficient estimates of linear regression models. Dependent variable: Post-Consumption Happ. is the happiness level with the chosen/allocated option after consumption occurred and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyNoChoice=1 if is in the No-Choice condition, and 0 if is in the Choice condition; dummyHighPOCs=1 if is in the High POCs condition, and 0 if is in the Low POCs or in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyMovie=1 if (post-consumption) happiness is reported for the Movie good category, and 0 if is reported for the Ice Cream good category. Standard errors in columns 1 and 2 are clustered at the individual level. Standard errors in parentheses, * p < 0.1; ** p < 0.05; *** p < 0.01.*

[59] Controlling for gender and age also does not affect the results.

In summary, we do not find support for hypotheses H1 and H2 when considering post-consumption utility. POCs therefore do not seem to significantly affect individuals' utility after consumption took place.

4.5 Discussion

The aim of the present study was to investigate whether the rejection of an alternative option implied by decision making negatively affects post-choice and post-consumption utility. We also tested if the size of this effect is positively related to the attractiveness of the forgone option. When considering post-*choice* utility, we found that conditional on obtaining the best option, individuals were worse off when this outcome resulted from a choice rather than from a situation without any choice. However, contrary to what we hypothesized, this effect was not especially stronger when the rejected option was an attractive one. When testing our hypotheses using post-*consumption* utility, we showed that rejecting an option, independently of its attractiveness, did not generate lower levels of utility after consumption took place.

In the following, we first discuss our findings in light of competing theories. We then mention some policy implications of our results.

4.5.1 Alternative explanations

Contrary to traditional economic theory, we find suggestive evidence that opportunity costs matter *after* a choice was made. However, further research is needed to properly rule out alternative phenomena that could potentially also explain our results. First, this effect may be driven by a spill-over of decision difficulty to the post-choice experience. The main idea behind decision difficulty is that making a choice (especially between closely ranked options, and among highly attractive options) may involve difficulties which might persist even after the choice is made, effectively reducing one's satisfaction. We cannot entirely rule out this possibility using the current paradigm. When controlling for a measure of

subjects' tendency to experience decision difficulty (captured by the variable *Choice Diff.* in Table C9, Appendix C)[60], regression results indicate that the subjects in the Choice group are still characterized by a lower reported post-choice happiness than those in the No-Choice group (see columns 1 and 2 of Table C9). However, our treatment effect becomes only weakly significant ($\hat{\beta}_{dummyChoice.}$ =-3.41, p-value<0.1). Therefore, further research would be needed to find out whether POCs can still explain our findings when the phenomenon of decision difficulty has been properly controlled for.

Another phenomenon that could explain our results is anticipated regret. We described in section 4.2.2 and in the presentation of the results that the feeling of regret cannot serve as an explanation of our findings for post-choice utility since the resolution of uncertainty regarding the utility of the chosen option has not yet taken place. However, while regret per-se cannot occur in the pre-consumption stage, *anticipated* regret could serve as an alternative explanation. Anticipated regret is basically the fear of experiencing regret over one's choice in the future, and it has been shown to play an important role in decision making (Mellers et al., 1997). Anticipated regret can hardly be ruled out as an alternative explanation to our pre-consumption stage. Further research would therefore be required to confirm the effects of POCs on post-choice utility.

4.5.2 Relevance

While additional research is needed to properly rule out alternative explanations to our results, we argue that POCs are an important overlooked precursor to satisfaction, and as such, a significant contribution to fields such as consumer research and judgement and decision-making. POCs might have substantial implications for other prominent phenomena such as consumer preference for a no-choice option (Dhar, 1997), the "keeping doors open"

[60] *Choice Diff.* is obtained by aggregating subjects' answers to the Decision Difficulty scale mentioned in section 4.3.2.

effect (Shin and Ariely, 2004), the choice overload phenomenon (Iyengar and Lepper, 2000)[61], and might perhaps even be a fundamental precursor of regret, along with its comparative component and self-blame (Connolly and Zeelenberg, 2002).

A deeper understanding of the relevance of POCs might be crucial for alleviating their detrimental effects on citizens' welfare by providing potentially valuable insights for public policy. People may be better off when having to respect rules imposed by the government than facing freedom of choice. For instance, a fixed duration of yearly holidays imposed by the policymakers or Sunday Blue Laws may reduce psychological opportunity costs and thus increase people's well-being. However, whether the benefit of reduced opportunity cost can outweigh the potential detrimental effect of less freedom to act upon one's preferences is still a matter of controversy. Refined evidence about the relevance of POCs is therefore necessary to further examine the simple standard economic assumption that freedom of choice necessarily leads to a higher consumers' well-being.

4.6 Conclusion

While economic theory assumes that opportunity costs only matter in the decision-making process, we argue that they remain relevant *after* choices have been made. More specifically, we assess the effects of opportunity costs, created by the forgone alternative, on the utility of the chosen option. We denote the extent to which opportunity costs matter for post-choice utility as *Psychological Opportunity Costs* (POCs), and argue that POCs are distinct from phenomena such as regret or disappointment.

[61] Choice overload deals with the paradoxical effect of being less satisfied with a choice made from a larger, rather than a smaller choice set. POCs could explain this phenomenon: while a bigger set size allows for a closer match with one's preferences, it simultaneously may increase the probability of having an attractive alternative in the set of forgone options and therefore of facing high opportunity costs. POCs might thus work as a mediator of the choice overload effect.

We conduct an incentivized laboratory experiment in which we randomize the opportunity costs that people face, i.e., the availability and desirability of an inferior option, and test their effect on utility, approximated by measures of expected and experienced well-being. More precisely, participants provide their happiness ratings with the (eventual) reception of either their favorite good (*No-Choice* condition), or with their choice out of two options – their favorite good and another one (*Choice* condition). In the latter case, we exogenously vary the attractiveness of the second option to generate either high or low opportunity costs: in the *High POCs* treatment, the alternative option is subjects' second preferred one, while in the *Low POCs* treatment, the other option is their least preferred one from a set of 6 items.

Overall, we find mixed support for our hypotheses. First, we show that post-*choice* satisfaction is significantly lower in the presence of opportunity costs, i.e., post-choice happiness ratings are lower in the Choice condition than in the No-Choice condition. However, we find that, on average, subjects in the High POCs condition do not report a significantly lower well-being right after the choice, than subjects in the Low POCs group. These results suggest that the simple rejection of an alternative option, independently of its attractiveness, can negatively affect one's satisfaction with one's choice. This finding is surprising as we would expect individuals to experience a bigger psychological loss when forgoing their 2nd best rather than their 6th best option. A reason may be that the rejected alternative in the Low POCs condition was still too desirable, or that its desirability was not sufficiently distant from the forgone option in the High POCs treatment. Another explanation may be that subjects did not sufficiently care about the outcome of their decision and therefore did not pay much attention to the unchosen option.

When analyzing post-*consumption* happiness ratings, we find that subjects in the Choice condition do not experience a lower utility than those in the No-Choice group. There

is also no difference in utility between participants who faced high or low POCs after consumption occurred. Therefore, POCs do not seem to matter anymore for post-consumption utility. The reasoning here may be that opportunity costs are only detrimental for people's well-being when they are salient in people's mind which may be the case until the last moment of consumption but not afterwards. However, this hypothetical explanation would need to be tested empirically.

Given the increasing amount of possibilities and decisions daily encountered by individuals, it is fundamental to better understand the factors influencing people's satisfaction with their choices. This is what we intended to do with the current study. We argue that POCs is a potentially relevant overlooked phenomenon that may negatively affect decision makers' happiness and encourage any future research that would further study this new concept and disentangle our findings from potentially alternative explanations.

Chapter 5 – Immoral Labor Markets

Joint with Florian Schneider and Roberto A. Weber[62]

5.1 Introduction

Corporate scandals regularly bring to light immoral work practices. For example, recent scandals in the financial industry reveal bankers' intentional sale of toxic assets to unsuspecting clients (US Department of Justice, 2016), manipulation of the London Interbank Offered Rate (LIBOR) (Abrantes-Metz et al., 2012) and aiding of tax evasion (Hill, 2012). Immoral practices also extend to other sectors. Tobacco company employees have long been accused of regularly engaging in unethical tactics such as misleading the public about the harmful effects of smoking (Heath, 2016) and developing marketing strategies to attract smokers as young as 13 (Bates and Rowell, 1998). Concerns about the potential immoral nature of professional activities are often discussed with respect to other sectors, such as the manufacture of weapons, the provision of high-interest credit to unsophisticated consumers and the "off-label" marketing of opioids.

While such immoral behaviors may sometimes result from the actions of a few rogue employees, there nevertheless exists a widespread impression that some jobs—e.g., marketing tobacco products, manufacturing weapons, engaging in predatory lending—likely involve inherently immoral acts by their nature. Conventional wisdom further posits that workers performing these kinds of jobs are compensated for their unethical conduct—a form of compensating differential for the aversive nature of acting immorally—and that such jobs attract those individuals who experience the least displeasure from committing immoral acts. Under this view, immoral work may represent an instance of standard arguments

[62] Please cite as Brun, Fanny, Schneider, Florian and Roberto A. Weber (2018) "Immoral Labor Markets," Working Paper.

regarding sorting and compensating differentials—in the same way as for other aspects of work that people may find heterogeneously unpleasant, such as risk of physical harm (Rosen, 1986).

However, despite the intuitive appeal of such a connection, there exists little empirical evidence that links heterogeneity in the willingness to perform immoral jobs to resulting differential labor market outcomes. Perhaps the closest evidence comes from studies documenting positive relationships between the immorality of work and the wages obtained by workers in those firms or industries (Frank, 1996; Moffatt and Peters, 2004; Benedict, McClough and McClough, 2006). For example, Frank (1996) used data from a Cornell University employment survey that included graduates' occupations, reported salaries and employers. He then asked students in a business ethics class to rate the "social responsibility" of the most common occupations and employers of Cornell graduates. A regression controlling for other observable characteristics—such as a student's major, grades and gender—reveals a premium for occupations and employers that are rated as socially irresponsible. Frank complements this evidence with other examples of cases in which professions that are otherwise similar but differ in their morality appear to produce different wages, such as anecdotal accounts that expert testimony on behalf of tobacco companies is "compensated handsomely."

Frank's evidence is consistent with the notion that heterogeneous concerns for morality and sorting result in differential wages across occupations and industries. However, there remain important gaps in documenting that the relationships observed by Frank are really the result of sorting, heterogeneous moral preferences and compensating differentials. Most obviously, correlational evidence between wages and the perceived immorality of work might result from other unobserved characteristics of workers and the work activities. For instance, Moffatt and Peters (2004) document a wage premium for prostitution, which they attribute

to a compensating differential; but it is unclear whether the compensation is for the perceived immorality of the work or other aversive aspects of the job (Edlund and Korn, 2002).[63] Moreover, such studies fail to measure a critical element—the identification of heterogeneous concerns for morality as a key driver of the relationship. While there is some evidence that people in some industries and professions exhibit lower concerns for morality (Sjöberg and Engelberg, 2009), no study identifies that these concerns for morality drive differential selection into different kinds of work rather than the relationship being, perhaps, the other way around (Frank, Gilovich and Regan, 1993; Cohn, Fehr and Maréchal, 2014).

In this paper, we provide novel evidence supporting the above relationships between the immorality of some kinds of work, worker's concerns for morality and the kinds of outcomes observed in "immoral labor markets." We do so with a combination of laboratory experiments, surveys and labor-market data, in which we obtain measures of individuals' concerns for morality and relate these to variation in the morality of work. The control provided by laboratory experiments allows us to observe the kinds of outcomes that arise as the nature of work changes *only* in the extent to which it is immoral. Complementary evidence from surveys and labor market outcomes allows us to obtain insights into relationships between individuals' concerns for morality and the morality of different firms and industries in real markets.

Our work is focused on two specific hypotheses that arise from a simple theoretical analysis of how individuals' heterogeneous aversion to acting immorally will interact with jobs that vary in the immorality of the work they require. We do not attempt to provide a novel theoretical contribution, but rather use simple economic analysis as a framework for formalizing intuitions and guiding our empirical research. The two main hypotheses simply

[63] Related work in finance (e.g., Hong and Kacperczyk, 2009; Colonnello, Curatola and Gioffré, 2017) demonstrates that investing in firms that engage in immoral activities ("sin stocks") yields higher returns. However, other aspects, such as risk, may also differ between these types of investments.

reflect the widely held perceptions that we note above: first, that more immoral work should yield higher wages (and lower quantities) than work that is not immoral and, second, that immoral work should attract those workers who care the least about morality. We then use our laboratory experiments, survey evidence and labor-market data to provide tests of these hypotheses. A critical novel contribution of our work is to directly relate differences in labor market outcomes to heterogeneity in individual concerns for morality.

Our results provide support for the above two hypotheses in both laboratory and field settings. First, we show that immoral work commands a wage premium over comparable work that is not immoral. We show this using Swiss labor market data, where we attempt to control for observable worker and industry characteristics and find that industries that are perceived to be immoral yield higher average wages. Moreover, in our laboratory labor markets—which vary by treatment only whether being employed requires doing something immoral, while holding other aspects of the job constant—we find that wages are persistently higher for immoral work. This wage premium is large and does not decrease with experience in the market.

Second, we provide evidence of sorting by immoral types into immoral work, both in the laboratory and in the field. We obtain two measures of individuals' aversion to acting immorally, one from a survey and one from a behavioral task. In our laboratory labor markets, immoral types are hired much more often only when work involves doing something immoral—i.e., there is no difference when the labor market does not involve immoral work. In our survey data, immoral types report a significantly greater willingness to work in firms and industries that are perceived to be immoral. While this result does not necessarily imply that they will actually end up working in these industries—though, importantly, Wiswall and Zafar (2018) provide evidence that such stated preferences are predictive of ultimate employment—it demonstrates that stated work preferences are

consistent with the sorting process in our second hypothesis. Moreover, our finding indicates that the (perceived) immorality of industries and firms might be identifiable by a revealed (or stated) preference approach—that is, immoral firms or industries are those that are relatively more attractive for immoral types as opportunities for employment.

Our work provides the first evidence documenting a differential willingness of heterogeneous moral types to work in jobs and industries that vary in their immorality. Importantly, this sorting persists with experience in our laboratory experiment and extends to stated preferences regarding real labor-market outcomes. We also connect this sorting to wage premiums for immoral work. In our field data this connection is correlational and thus subject to caution in drawing interpretations from the relationship, but in our laboratory we demonstrate a causal relationship by using a design that varies by treatment only whether the work is immoral.

Our study also relates to research on effort and sorting across different kinds of work by "mission-oriented" types (Besley and Ghatak, 2005; Ariely, Bracha and Meier, 2009; Fehrler and Kosfeld, 2014; Carpenter and Gong, 2016; Cassar, forthcoming; Cassar and Armouti-Hanse, 2018). There are similarities between this literature and our work and one interpretation of our research is that we study the opposite type of preferences—anti-social rather than pro-social. However, we believe that there is something inherently different, and more general, about a desire not to do immoral things than a positive attitude toward a particular kind of mission. Moreover, the connections between carefully identified preferences, outcomes in laboratory labor markets, and real-world labor market outcomes in our work are novel even with respect to this literature.

The rest of our paper proceeds as follows. The next section presents evidence from Swiss labor market data that demonstrates a positive correlation between the immorality of work and wages. While similar results have been observed previously (e.g., Frank, 1996;

Benedict et al., 2006), our analysis involves a larger sample representative of a more diverse set of workers and jobs. We then present a simple theoretical framework that we use to motivate this and other relationships that one would expect to see in data under heterogeneity in moral concerns and work characteristics. Section 5.4 presents the design of our laboratory experiment and survey studies, while section 5.5 presents the results. We conclude in section 5.6 by discussing implications of our findings.

5.2 Suggestive correlational evidence from the Swiss Labor Force Survey

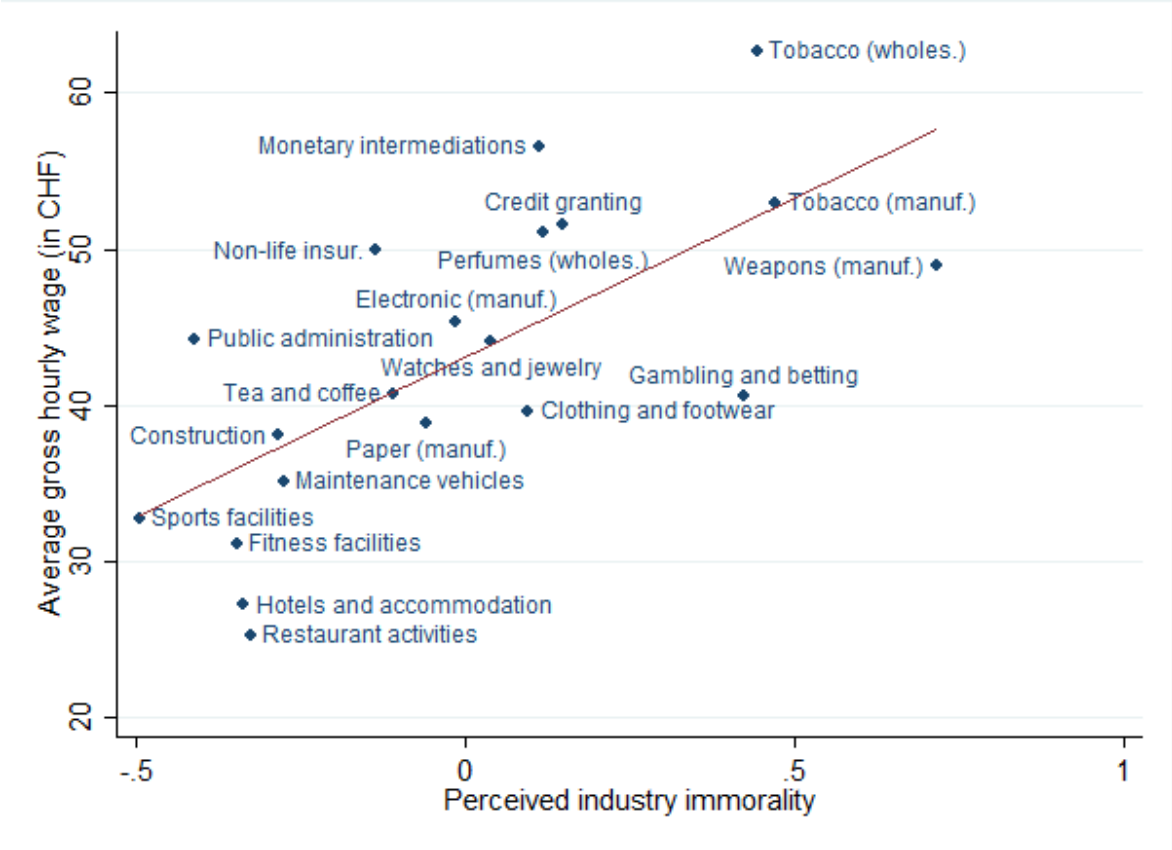
As an initial step, we first explore whether individuals working in "immoral" industries receive a wage premium. For this, we use data from the Swiss Labor Force Survey (SLSF)—a representative sample of the Swiss labor force compiled by the Swiss Federal Statistical Office—to explore the relationship between the perceived immorality of work and the portion of wages that cannot be easily explained by observable worker characteristics.

In this exercise, we first identified industries that we jointly perceived as involving high degrees of immoral work activities; we did so before looking at any data, including wages, from these industries.[64] This yielded six "immoral" industries: gambling and betting activities, monetary intermediations, credit granting, manufacture of tobacco products, wholesale of tobacco products and manufacture of weapons and ammunitions. We then chose comparison industries from within the same industrial branch with similar distributions of education levels, as well as a few additional industries representing large shares of employment in Switzerland. We did not look at wages when we were selecting the comparison industries. To independently validate our judgments that these sets of industries

[64] Specifically, we started with the complete list of industries listed in the dataset. Each of the three authors went through the list and indicated any industries that he or she thought had a significant immoral component. We selected those industries for which all three authors agreed.

differed in the extent to which they are moral, or at least perceived as such, we asked an independent sample of 177 students to rate each industry on a scale ranging from "very moral" (1) to "very immoral" (9) and re-scaled the responses to lie on the -1 to 1 interval. (These data were collected as part of our survey studies, which we describe in more detail

Figure 13: Correlation between wages and perceived industry immorality



Source: Weighted data from the SLFS, years: 2010, 2012, 2014 and 2016 (wage) and our own survey (perceived immorality). Notes: Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Real gross hourly wage in 2010 CHF. Cross-section with N = 18,514. Number of observations per industry: Credit granting=44, Monetary intermediations=4146, Gambling and betting activities=91, Manufacture of tobacco products=88, Manufacture of weapons and ammunitions=57, Wholesale of tobacco products=51, Non-life insurance=1484, Organization and operation of sport facilities=288, Processing of tea and coffee=75, Manufacture of electronic components=656, Wholesale of perfume and cosmetics=200, Wholesale of clothing and footwear=166, Wholesale of watches and jewelry =91, Manufacture of paper and paperboard=54, Maintenance and repair of motor vehicles=1508, Construction of buildings=2022, Restaurant and mobile food activities=3137, Hotels and similar accommodation=1514, Fitness facilities=227, General public administration activities=2615. For credit granting: extrapolation based on less than 50 observations; this result must be interpreted with great caution.

Table 15: Relationship between wages and perceived industry immorality

Dependent variable: ln of real gross hourly wage (in 2010 CHF)			
	(1)	**(2)**	**(3)**
Perceived industry immorality	0.936***	0.681***	0.372***
	(37.60)	(29.73)	(11.13)
Age		0.006***	0.005***
		(11.76)	(11.03)
Male		0.200***	0.181***
		(16.19)	(15.30)
Married		0.037***	0.027**
		(3.83)	(2.89)
Education high		0.510***	0.390***
		(28.39)	(19.25)
Education middle		0.187***	0.132***
		(11.14)	(7.45)
Swiss		0.063***	-0.022*
		(6.16)	(-2.18)
Tenure		0.005***	0.003***
		(7.92)	(5.52)
Full-time equivalent		-0.031	-0.054*
		(-1.13)	(-1.99)
Board member		0.031*	0.091***
		(2.08)	(6.68)
Constant	3.767***	2.990***	3.128***
	(594.96)	(90.89)	(86.59)
N	18514	18514	18514
Adjusted R^2	0.133	0.330	0.394
Industrial branch FE	No	No	Yes

*Source: Weighed data from the SLFS, years: 2010, 2012, 2014 and 2016 (wage) and our own survey (Perceived immorality). Notes: Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: Male in {0, 1}, Married in {0, 1}, Education high: higher vocational education and training or university/college, Education middle: apprenticeship, full-time vocational school, matura or pedagogical training, Education low (reference category): compulsory schooling or pre-vocational education, Swiss in {0, 1}, Tenure = number of years in the firm, Full-time equivalent = (working hours /42), set to 1 for working hours >= 42, Board member in {0, 1}. Model (3) controls for industrial branch fixed effects (manufacturing; wholesale and retail trade; repair of motor vehicle and motorcycles; financial and insurance activities; accommodation and food service activities; construction; public administration, defense and compulsory social security; arts, entertainment and recreation). Robust standard errors, t-statistics in parentheses; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

in section 5.4.) The mean ratings for each industry are shown on the horizontal axis of Figure 13. They confirm that our initial judgments with respect to the perceived immorality of certain industries are shared by the student sample.

The vertical axis of Figure 13 plots the mean real gross hourly wage (in 2010 Swiss Francs) in each industry. These data are the reported hourly wage of employees surveyed as part of a national representative panel. We use cross-sectional data from the 2010, 2012, 2014 and 2016 waves. The strong positive relationship supports the hypothesis that work in less moral industries yields a wage premium. The size of the differences is also substantial.

Of course, the relationship in Figure 13 ignores the potential role of individual worker characteristics, which may vary across industries and may explain some of the differences in wages. To partially address this concern, Table 15 reports regressions of the hourly wage reported by individuals in different industries on the perceived immorality of each industry, along with several additional control variables (Appendix Figure D1 provides a scatter plot of the regression residuals with respect to perceived immorality). Model 1 displays the results of a simple regression of the natural logarithm of real gross hourly wages on the perceived industry's immorality, supporting the positive relationship in Figure 13. Model 2 adds observable worker characteristics, while Model 3 additionally includes indicator variables for the industrial branch to which an industry belongs. While the addition of these controls lowers the magnitude of the industry immorality coefficient, the immorality premium remains substantively large and statistically significant: according to Model 3, individuals working in a totally immoral industry (i.e. Perceived immorality = 1) have (geometric) mean hourly earnings approximately 45 percent higher than people working in a neutral industry (i.e. Perceived immorality = 0).[65]

[65] We obtain this number by doing the following calculation: $e^{0.372} - 1 \approx 0.45$. Summary statistics of all variables used in the regressions and all of the industries are provided in Appendix Table B.10.

The above analysis is consistent with the notion that workers are being compensated for the immoral nature of some kinds of work. However, the correlational aspect of the relationship leaves open the question of whether unobserved characteristics of the industries may explain the relationship in Figure 13. Moreover, the above analysis does not tell us whether the workers employed in these industries differ in their concerns for morality. To explore this question much more carefully—controlling for unobservable aspects of the work and making a clearer connection to individual workers' heterogeneous concerns for morality—we turn to the theoretical framework guiding our survey and laboratory studies.

5.3 Model

In this section, we introduce a simple stylized model of labor markets with varying degrees of immorality of jobs and heterogeneity in concern for morality among workers. The purpose of the model is to formalize the intuition about how varying immorality of jobs interacts with workers' heterogeneous concerns for morality. We use the theoretical results to guide our investigation of immoral labor markets, in particular for designing the experiment and analyzing the data. Our framework is a simplification of the theoretical literature on compensating wage differential (see e.g. Rosen, 1986).[66] We do not seek to expand this literature, but rather to apply it to a context in which the relevant characteristic is morality.

Workers are hired by firms to do a job $j \in J$. The job might involve doing immoral work. The immorality of the job is measured by a function $I: J \to \mathbb{R}_{\geq 0}$, where $I(j') >$

[66] Unlike most models of compensating wage differentials, we do not have multiple labor markets, but only one, along with a fixed outside option. In the experiment, we also assign subjects to one labor market. This abstraction simplifies both the theory and the experiment. However, we show in Appendix D that our model allows for an interpretation with two jobs, an immoral job and a neutral job. Our results therefore also apply to such a context.

$I(j)$ means that job j' is more immoral than job j, and $I(j) = 0$ means that the job j involves no immoral acts. The set of immoral jobs is $J^{IM} = \{j \in J: I(j) > 0\}$.

Labor demand is represented by an interval of firms, $k \in [0,1]$. Each firm hires workers for the same job j. Firms' behavior is given by the labor demand function, $D: \mathbb{R} \times J \to [0,1]$, where D is continuous in w, strictly decreasing in w on $[0, \infty)$, with $lim_{w \to \infty} D(w,j) = 0$ and $D(w,j) = 1$ for $w \leq 0$. In addition, we assume that an increase in immorality of the job does not decrease profitability of labor, that is, $I(j') > I(j)$ implies $D(w,j') \geq D(w,j)$ for all $w \in \mathbb{R}$.[67]

Labor supply consists of an interval of workers, $i \in [0,1]$. Each worker has reservation utility $\underline{u} \in \mathbb{R}_{\geq 0}$, and the utility of accepting job j of a worker of type i is given by:

$$U_i(accept|j,w) = w - c - I(j) * \theta_i,$$

where $c \in \mathbb{R}_{\geq 0}$ is the worker's cost of effort. The parameter $\theta_i \in \mathbb{R}_{\geq 0}$ measures how much the worker cares about the immorality of the job and is distributed according to a cumulative density function $F \in \mathcal{F}_\theta$. For all $F \in \mathcal{F}_\theta$, F is continuous, strictly increasing on $[0, \infty)$, and with $F(0) = 0$.[68] The utility of a worker of type i is then given by $U_i(j,w) = max\{\underline{u}, U_i(accept|j,w)\}$. Workers' behavior determines the labor supply, $S: \mathbb{R} \times J \to [0,1]$. If $j \in J^{IM}$, every worker with $\theta_i \leq \frac{w-\underline{u}-c}{I(j)}$ accepts the job. Labor supply is therefore $S(w,j) = F(\frac{w-\underline{u}-c}{I(j)})$.[69]

[67] If an increase in immorality would decrease profitability, there would be no incentives for firms to operate in immoral industries. Heidhues, et al., (2017) provide a basis for why deceptively marketed socially harmful products may be more profitable in the presence of naïve consumers. In Appendix D, we provide a behavioral foundation for labor demand. In our experiment, we vary the immorality of the job, but fix labor demand, that is, $D(w,j') = D(w,j)$ for all w and all $j, j' \in J$.

68 Note that $F(0) = 0$ implies that no worker likes to do immoral jobs (see also Rosen (1986, p. 645).

[69] The assumptions on F (together with the properties of a cdf) imply that S is continuous and strictly increasing in w on $[\underline{u} + c, \infty)$, $\lim_{w \to \infty} S(w,j) = 1$, and $S(w,j) = 0$ for all $w \leq u + c$.

Using this framework, we can now consider the equilibrium properties of this type of market. The equilibrium wage, $w^*(j)$, is implicitly defined by $S(w^*(j),j) - D(w^*(j),j) = 0$.[70] The following Lemma states that for every $j \in J$, $w^*(j)$ exists and is unique (all proofs are in Appendix D).

Lemma. *For all* $j \in J^{IM}$, $w^*(j)$ *exists, is unique and is in* $(\underline{u} + c, \infty)$. *For all* $j \in J \setminus J^{IM}$, $w^*(j) = \underline{u} + c$.

In the following, we derive four properties of labor markets with immoral jobs. While straightforward, these results yield predictions for our empirical work. In particular, the first two propositions provide the primary hypotheses that we test across all of our analysis. The third and fourth results are testable in our laboratory experiment.

Proposition 1. *For all* $j, j' \in J$ *with* $I(j) < I(j')$, $w^*(j) < w^*(j')$.

Proposition 1 shows that there is an immorality premium for immoral jobs: an increase in the immorality of a job decreases supply and therefore increases the equilibrium wage. Indeed, we have already demonstrated evidence supporting the wage relationship in this hypothesis in Figure 13, where variation on the horizontal axis corresponds to variation in $I(j)$.

Second, the immoral types, or to be precise, the types that care least about the immorality of a job ($\theta_i \leq \frac{w^*(j)-\underline{u}-c}{I(j)}$), sort into accepting immoral jobs, while those workers

70 Note that for $j \in J \setminus J^{IM}$, S is a correspondence. For this case, $w^*(j)$ is defined by $D(w^*(j),j) \in S(w^*(j),j)$. Moreover, $w^*(j)$ depends on F. When necessary (Proposition 4) we will make this explicit by writing $w^*(j,F)$ instead of $w^*(j)$.

more concerned with morality ($\theta_i > \frac{w^*(j)-\underline{u}-c}{I(j)}$), refuse to do the job for the equilibrium wage.[71] This is formally shown in Proposition 2.[72]

Proposition 2. *For all* $j \in J^{IM}$*, worker* i *is hired iff* $\theta_i \leq \frac{w^*(j)-\underline{u}-c}{I(j)} \equiv \underline{\theta}(j) \in \mathbb{R}_{>0}$.

Proposition 2 is critical to the notion that immorality wage premiums like those in Figure 13 are driven by those who find immoral work most distasteful opting out of jobs that require immoral behavior. This important relationship has, to our knowledge, not been previously tested.

Our next two predictions are less central to our purposes, but nevertheless provide some useful and testable insights into behavior in immoral labor markets.

Proposition 3. For all $j, j' \in J$ with $I(j) < I(j')$, there exists $\tilde{\theta}(j, j') \in \mathbb{R}_{>0}$ such that $U_i(j', w^*(j')) > U_i(j, w^*(j))$ iff $\theta_i < \tilde{\theta}(j, j')$.

Proposition 3 shows that immoral types profit from an increase in the immorality of work. More precisely, there are some types who are sufficiently unconcerned with morality who are hired in an immoral market and are overcompensated by the immorality premium.

Finally, Proposition 4 shows that moral behavior (refusing to do the immoral job) can have positive externalities for immoral types. If the distribution of types shifts towards more concern for morality—in the sense of stochastic dominance—then the supply at any given wage decreases, thereby increasing the equilibrium wage and the utility of those least

71 Note that this perfect sorting according to θ is an extreme, and admittedly unrealistic, case. If there is heterogeneity in the costs of effort, reservation utility or productivity, there is only partial sorting by θ. We do not incorporate more than one dimension of heterogeneity in our model to keep it simple. Heterogeneity in both productivity (earnings capacity) and risk preferences is investigated in Garen (1988) and Hwang, Reed and Hubbard (1992).

72 Note that for any $a \in \mathbb{R}_{>0}$, $F(a) > 0$ because $F(0) = 0$ and F is strictly increasing on $[0, \infty)$. Therefore $F(\underline{\theta}(j)) > 0$, implying that some workers are hired. This is also important for most of the other propositions.

concerned with morality. For instance, any completely immoral types (i.e., $\theta_i = 0$) will always benefit from a higher wage produced by increased moral concerns.

Definition. *For all* $F, G \in \mathcal{F}_\theta$, F *strong first-order stochastically dominates* G *iff* $F(x) < G(x)$ *for all* $x \in \mathbb{R}_{>0}$.

Proposition 4. *For all* $j \in J^{IM}$ *and* $F, G \in \mathcal{F}_\theta$ *with* F *strong first-order stochastically dominating* G, *there exists* $\hat{\theta}(j,F) \in \mathbb{R}_{>0}$ *such that* $U_i(j, w^*(j,F)) > U_i(j, w^*(j,G))$ *iff* $\theta_i < \hat{\theta}(j,F)$.

5.4 Study design

Our study uses a combination of surveys and a laboratory experiment to measure heterogeneity in concerns for morality (θ_i) and to relate this heterogeneity to outcomes as labor markets vary in the immoral nature of work ($I(j)$). Table 16 provides an overview of our design and how we obtain these measures.

Table 16: Overview of study design and measures

Workers (n = 240)		Independent raters / "Clients" (n = 177)
On-line survey	**Laboratory experiment (4-7 days later)**	**Survey study**
Concern for morality (survey-based: θ^{Sur})	Concern for morality (behavioral task: θ^{Exp})	Rated the immorality of firms and industries ($I(j)$)
Labor market expectations (industries, firms, wages)	Repeated labor market; subjects can be hired to give advice to "clients;" treatment varies job immorality ($I(j)$)	Role of passive "clients" for laboratory experiment
Personal characteristics (e.g., age, field of study, Big 5)		

The study uses two different samples. Our main focus is on the sample of *Workers*, who participate in our laboratory sessions. These subjects initially complete an on-line questionnaire that includes several questions designed to measure their concern for morality.

From these items, we construct an individual measure of concern for morality (θ^{Sur}). At this point, we also attempt to measure the subjects' expectations of their own future labor market outcomes, including the willingness to work for different firms and industries and expected future wages.

These same subjects then participate in a laboratory session roughly one week later. In this session, they perform two choice tasks. First, we elicit concern for morality using an incentivized behavioral task, adapted from Gneezy, Rockenbach and Serra-Garcia (2013), that creates a tradeoff between personal monetary gain and moral conduct. Workers then participate in a laboratory labor market for 15 periods, in which they submit reservation wages for being hired to perform work. Labor demand is simulated according to a fixed demand schedule.

The key feature of our laboratory labor market is that our treatment varies only the degree to which work is immoral, while holding constant all other job characteristics, including the specific actions workers take when employed. More precisely, subjects in our experiment are hired to provide written advice to a "client." We simply vary whether the recommendation is honest (in which case a choice is recommended that the client very likely would have made anyway, which earns the client some money and produces a donation to a UNICEF fund that provides malaria treatment for children) or dishonest and recommends the only choice that eliminates the above benefits.[73] The degree of immorality of jobs exogenously varies: workers are either assigned to a market with neutral jobs or a market with immoral jobs.

[73] There are several ways to design an immoral task in the laboratory. We build our task on bad advice giving and honesty. This has two advantages: first, misleading consumers is a realistic feature of many existing "immoral" jobs. Second, there is an extensive psychological and economic literature on the measurement of honesty and lying aversion (e.g. Gneezy, 2005; Gneezy et al., 2013) and having a reliable measure of concern for morality in the context of our market is important to test the predictions of our model.

Finally, we recruit a separate sample of individuals at public locations. These participants serve two functions. First, they act in the role of clients: they receive written recommendations from laboratory subjects and decide whether to follow them. From these choices, the clients accumulate money and influence the size of a donation to UNICEF. Second, these participants complete a survey in which they evaluate the extent to which various industries and firms are "moral" or "immoral." These are the ratings that we already used in Figure 13 and Table 15.

5.4.1 The on-line survey

Students who were recruited for laboratory sessions were asked to complete an on-line survey about 4-7 days before coming to the lab. For our purposes, there were two main components to the survey.

First, we included several multi-item scales intended to measure a broad concern for morality and moral acts. These were:

1) **Charity attitude index.** We used a 9-item scale developed by Brashear et al. (2000) in which participants rate statements regarding how positive and useful they perceive work done by charities and how important they perceive it is to help others in society.
2) **HEXACO-PI.** We administered 10 items from the short version of the HEXACO Personal Inventory (Ashton and Lee, 2009) related to the factor "Honesty-Humility"—consisting of the four traits, sincerity (3 items), fairness (3 items), greed avoidance (2 items) and modesty (2 items). Every item describes a thought that a moral or immoral person might have and participants indicate the extent to which each thought reflects their own opinions.
3) **Protected Values.** The Protected Values scale (Tanner et al., 2009; Gibson et al., 2013) measures an individual's position regarding values that can be seen as

inviolable, and not substitutable against money, and that are usually central to the person's identity. In our case, and following Gibson et al. (2013), we adapted the Protected Values to a situation where a financial adviser can give bad investment advice to a client for personal benefit. First, 5 items assess the morality of this behavior (*Protected value 1*); second, 4 items examine how truthfulness matters in such a situation (*Protected value 2*).

4) **Integrity and Work Ethics Test.** We used two items from a test designed to allow firms to measure the integrity of job applicants (*Work ethics 1, Work ethics 2*). In each item, participants read two fictitious dialogues between two characters with different opinions about a situation (e.g. calling in sick at work to enjoy a sunny day outside). Participants then rate with which character they agree more.

In each case, subjects expressed agreement or disagreement with statements on either a 5-point or 7-point Likert scale. The description of these survey scales are provided in Table D6 in the Appendix.

Second, we asked subjects several questions about their future labor-market expectations. Subjects were shown a list of 26 well-known companies in Switzerland and another list consisting of the industries in Figure 13. Both lists are available in Tables D12 and D13 in the Appendix. Participants rated their willingness to work for each firm and industry (1: *not at all willing;* 5: *very much willing*).[74] For firms, participant also had the option to chose "I don't know this organization" instead of rating the willingness to work. In addition, we asked subjects to provide estimates of their career plans; specifically, what work they expected to do after their studies and how much they expected to earn at the age of 40.

[74] Incidentally, the list of industries missed five industries for the five participants that participated in the first lab session. So we miss their data on willingness to work in these industries for these participants. Other than this, they completed the full questionnaire.

Finally, we asked subjects whether several non-profit organizations (including UNICEF) are worth supporting and collected several additional personal characteristics, including a short version of the Big Five questionnaire (Gosling et al., 2003), which identifies individuals' extraversion, agreeableness, neuroticism, openness, and conscientiousness, but is largely orthogonal to morality. We also elicited several demographic characteristics, such as age, gender and field of study.

5.4.2 The laboratory experiment

In the laboratory experiment, sessions consisted of 24 participants. During the experiment, subjects accumulated earnings in "points," which were converted to money at the rate of 20 points = 1 CHF ≈ 1 US Dollar.

We initially took a portrait photograph of each subject. Participants were asked to hold a neutral face while the picture was taken. We did so because our labor market made outcomes public—to reflect the fact that individuals' choices of whether to work in immoral industries are typically visible to their friends and neighbors in real life. Thus, we allowed each subject in a labor market to observe the employment and earnings outcomes of each of the other participants in their labor market. Importantly, this means that our behavioral measure of "immorality" confounds both internally-driven concerns for acting immorally and concerns for being perceived as immoral by others. Since real-world labor markets typically also confound both motives, we decided to include them in our study. Moreover, for the purposes of this experiment, we attempted to design an "immoral" act that is strong in the extent to which many individuals likely find it aversive. Thus, many of our design features are intended to strengthen the degree of immorality, real or perceived, in the underlying acts.

Another important feature of our experiment is that subjects had the ability to influence the amount of a donation to UNICEF; specifically, to a fund that provides

treatment to children with malaria. This donation allowed us to strengthen the moral component of behavior in the experiment and to extend the impact of subjects' choices beyond the experimenter and other participants in the study. At the beginning of the study, participants read an information sheet about the consequences of malaria and the need for treatments—we adapted this material from public materials available from UNICEF. Each session started with an amount of donations sufficient to provide several children with anti-malarial treatment, per UNICEF's estimate.[75] However, the actual final donation was determined by the behavior of participants in the session (and the choices of the clients). To make the moral component strong, we further framed choices as helping to "save children" from malaria.[76]

Once all participants entered the laboratory, we first measured their concern for morality in an incentivized decision task used to elicit general willingness to lie and to harm others for monetary gain. Next, participants played 15 periods of an experimental labor market. After the final period, we collected some additional individual-level measures, which we describe below. Finally, some subjects were randomly selected to perform the role of second-movers in the behavioral task that we conducted at the beginning of the session. Participants were then paid, in private. We now describe each of these steps in detail.

Behavioral measure of concern for morality ($\boldsymbol{\theta^{Exp}}$)

Participants first played an incentivized game that measures their willingness to lie for personal gain while causing harm to others. The task builds on a game by Gneezy et al.

[75] To further strengthen the moral component, we framed the donation as having been generated by a third party's blood donation. Specifically, prior to the laboratory session, we approached individuals who had just donated their blood as part of a donation campaign. We asked them whether they would agree that the University of Zurich make a donation to UNICEF as a complement to their blood donation. We used these donations to frame the origin of the initial endowment.

[76] We borrowed terminology from UNICEF's website, which states that the anti-malaria treatment saves children from the deadly effect of malaria.

(2013), and modifies it such that it mimics the consequences of a lie in the immoral treatment in our experimental labor market.

In the game, Participant A privately observes a computerized die roll and sends a message claiming the observed number to Participant B. Participant A may claim that the assigned number r is either "1", "2", "3", "4", "5" or "6" regardless of the actual number assigned. Participant A receives 100 + 20 * r points, which means that she has an incentive to lie if r is less than 6. Participant B then decides whether "to follow" or "not to follow" the message sent by Participant A. If he does not follow the message, he receives 30 points and the donations to UNICEF are unaffected. If he follows the message and Participant A truthfully reported the observed number, Participant B earns 100 points and the initial donation to UNICEF is increased by one additional anti-malarial treatment. However, if Participant B follows and Participant A lied, Participant B does not earn any points and the donation to UNICEF is reduced by one additional anti-malaria treatment.

Every participant initially plays the role of Participant A. We use the strategy method to elicit Participant A's message for every possible die roll. This allows us to classify all Participant A's by their strategies in the game. At the end of the experiment, 5 of the 24 participants in the session have their role changed from Participant A to Participant B. These Participants B are then matched with five of the remaining Participants A and decide whether or not to follow the corresponding message. All participants whose role is not switched—who remain as Participant A—are paid based solely on their choice as Participant A, independently of whether or not they are matched with a Participant B.[77]

Participants were informed that, at the very end of the session and after all choices had been made, their decisions as Participant A would be publicly displayed to 5 other

[77] This implies that Participant As (whose role was not switched) received their own payment with certainty, their decision, however, only had consequences for Participant Bs (and UNICEF) with a probability of 26.3 percent. This reflects a feature of our experimental labor market.

individuals, along with their portrait photograph. This was all explained clearly and publicly at the beginning of the experiment.

Market experiment

In the labor market, participants play the role of workers competing to be hired by automated firms. Before interacting in the market, instructions about the labor market are distributed to participants and read aloud using a voice recording. Then, the participants answer a list of understanding questions about the market, including how prices and quantities are determined. Only after the above instruction about market procedures, subjects receive information about the nature of the job. This ensures that subjects in both treatments interpret the market instructions in the same manner. Their understanding of these new instructions is again tested and corrected.

The job. In both markets, workers have the opportunity to be hired as an "advisor" whose job is to give advice to another uninvolved participant outside the laboratory, the "client." Specifically, the advisor has to write a recommendation to a client to choose one among ten choice options (labeled by the letters A through J). The client receives this recommendation, and has to choose one of the ten options. The client only knows that the option he selects determines his financial reward for completing a survey and influences an initial amount of money donated to UNICEF, but does not know the consequences of any specific option. Importantly, however, the client knows that the advisor had complete payoff information at the time of writing the recommendation. The client is free to choose the option that he is recommended or to choose any other option.

The payoffs associated with each option are indicated in Table 17. Nine options increase the reward of the client by 1 CHF (≈ 1 US Dollar) and increase the donation to UNICEF by an amount estimated to correspond to the anti-malarial treatment of one child.

However, one of the 10 options—in this case, option D—gives 0 CHF to the client, and reduces the donation to UNICEF.

Table 17: Options available to the "client"

	A	B	C	D	E	F	G	H	I	J
Additional number of children who receive the anti-malarial treatment	1	1	1	-1	1	1	1	1	1	1
Financial reward for client	1 CHF	1 CHF	1 CHF	0 CHF	1 CHF	1 CHF	1 CHF	1 CHF	1 CHF	1 CHF

Our two treatments vary only in the recommendation that the advisor is hired to make to the client. In the neutral treatment, the job of the advisor is to recommend one of the options that is beneficial to the client and to UNICEF (e.g., option G in Table 17). Note that a client is very likely to make such a choice independently of any advice. In the immoral treatment, the job is to recommend the single option with negative consequences (option D). In both cases, the advisor makes a recommendation by completing a form stating that, "option [G] will save the highest number of children" and "will give you the highest financial reward."[78] By recommending option D, the advisor increases the chance that the client selects the single option that will not increase his earnings and that will reduce the donation to UNICEF.[79] We vary the letter of the bad (neutral) option across immoral (neutral) laboratory sessions.

The market. Participants are randomly allocated to markets consisting of 6 workers who compete to be hired by 6 automated firms. Each worker can provide up to two units of labor—one at a low cost (50 points = CHF 2.50) and one at a high cost (110 points =

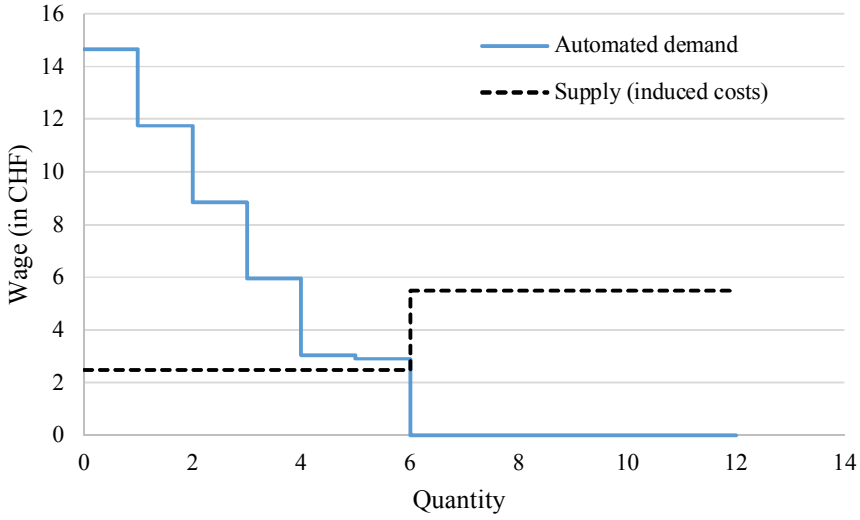
[78] Specifically, the advisor has to complete the following recommendation with the option's name (G) and his initials: „I, [advisor's initials], have reviewed your possible choices and I recommend that you select the option [G]. Following my advice will save the highest number of children and will give you the highest financial reward. Your advisor: [advisor's initials]"

[79] Subsequently, 84% of all recommendations produced in the laboratory were followed by clients.

CHF 5.50). The induced costs are the same for all participants, which the instructions clearly explain.

Each worker decides whether or not to participate in the labor market. In the former case, she (privately) provides two wage requests for each of the possible units of labor she can provide. Workers may only submit wage requests that are at least as high as the corresponding cost of providing that job. Each firm can hire up to one unit of labor per period. Firms are identical except for the wage that they offer to the workers. Figure 14 displays the automated demand for labor as well as the induced costs of labor supply. The workers have no information about the shape of the automated demand.

Figure 14: The automated demand and the induced costs of the labor supply



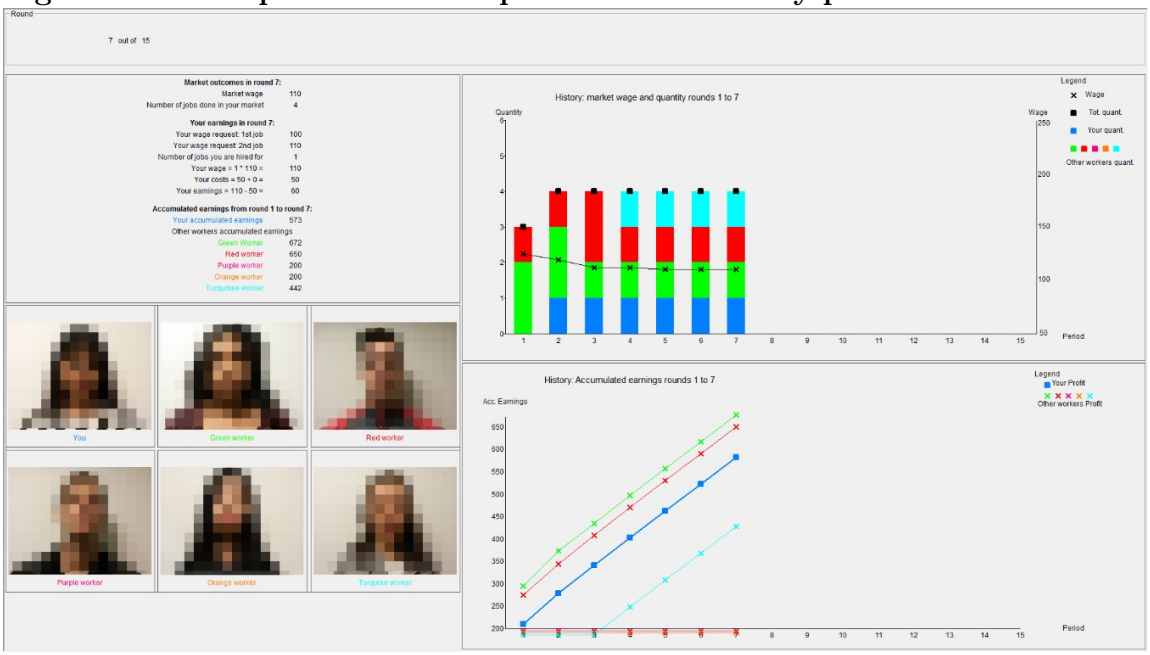
We use a uniform-price sealed-offer auction as the market mechanism, as this provides several advantages. First, Smith et al. (1982) show that this market converges to the equilibrium prediction. Second, and more importantly, this mechanism allows us to automate the labor demand (see also Sausgruber and Tyran, 2011) and therefore to keep the demand constant between the neutral and the immoral treatments. Once all six workers have submitted their wage requests, the computer ranks them from lowest to highest and compares the workers' wage requests to the firms' wage offers, ranked from highest to

lowest. The market wage is then the lowest number between two potential candidates: (i) the last wage offer that is higher than the wage request with the same rank, and (ii) the first wage request that is higher than the wage offer with the same rank. This mechanism clears the market in the sense that, for the market wage, labor supply equals labor demand and all workers with wage requests below the market wage are hired. If the firm hired no worker, the client receives no recommendation. This implies that in the immoral treatment, if a participant is not willing to do the job for the market wage, the number of clients who receive bad advices (weakly) decreases.

The market is repeated for a total of 15 market periods. The composition as well as the type (immoral or neutral) of each market is fixed over all periods. At the end of each period, the computer reports the market wage, displays the picture of every worker in the market and summarizes information regarding each workers' behavior and financial status across all periods (see Figure 15). Specifically, subjects observe employment outcomes, wages and cumulative earnings for themselves and for all other workers in their market over periods, and can connect these to the other workers' identities through the photograph. After observing outcomes, those participants who where hired in a period completed the paper forms with the recommendations—they wrote their own initials and the appropriate letter (e.g., "G" or "D" in the earlier example).[80]

[80] Subjects were informed that each firm had a 0.25 probability of actually having a client in each period, which was independent of whether or not the firm hired a worker. If the firm did not have a client, then the worker's recommendation would be unused, although the worker would still complete the recommendation and receive the market wage in that period. However, subjects did not know at the time of submitting wage requests or completing the forms at the end of a period whether or not there would actually be a client for this period. They were informed, at the end of the experiment, for which periods any recommendations they wrote would actually be distributed. We did this to lower the number of clients we would have to recruit as part of the follow-up survey. This represents, for instance, a case in which a worker is hired to prepare promotional materials for a harmful product, which may or may not ultimately be used in a marketing campaign. This procedure implies that writing a recommendation has only consequences with a probability of 25 percent.

Figure 15: Example of feedback provided after every period



Additional measures. At the conclusion of the laboratory session, we collected several additional individual-level measures. First, we measured participants' affect levels—i.e., pleasure, arousal, and dominance—using the self-assessment manikin (Bradley and Lang, 1994). Next, we asked participants whether they thought the clients would or would not follow recommendations; this belief was not incentivized. The participants were then prompted to answer several questions about the reasons underlying their market behavior. Finally, we measured subjects' concerns for social image using the public self-consciousness scale by Leary et al. (2015), in which participants rate seven short descriptions of behaviors by people who care or do not care about their self-image, on a scale from 1 (not like me at all) to 4 (a lot like me).

Procedural details. All sessions took place at the Decision Sciences Laboratory (DeSciL) at the Federal Institute of Technology in Zurich (ETH) in February, April and May 2017. We recruited participants using hroot (Bock et al., 2014) from the joint subject pool of the University of Zurich and the ETH. Every session consisted of 24 participants,

who are only accepted at the session if they had completed the on-line survey.[81] We conducted ten sessions, resulting in a total of 240 participants, allocated to 28 immoral markets and 12 neutral markets. The on-line questionnaire was implemented with the Qualtrics software (Qualtrics, Provo, UT) and the laboratory experiment with zTree (Fischbacher, 2007).

All instructions were delivered both on paper and with pre-recorded audio files. Instructions and materials are available in the Online-Appendix.[82]

5.4.3 Survey study

We subsequently recruited a different sample of students on the campus of the University of Zurich and the ETH (N=177). We invited student passersby to participate in a brief choice experiment in which they could earn money and generate a donation for UNICEF aimed at providing treatments for children infected by malaria. They were told that they would earn CHF 2 plus possibly some additional money for a 5-minute study. These subjects performed two functions.

First, they served the role of "clients" for the recommendations from the laboratory labor market. Each participant made up to six decisions by choosing one of the ten letters between A and J. They knew that these decisions influenced their own earnings and also possibly the amount of donations to UNICEF. Each decision had the payoff structure in Table 17, but we varied which letter corresponded to the bad option. Clients received a mixture of recommendations with good advice, bad advice or no advice (corresponding to the case in which a firm was not able to hire a worker). Clients were only informed of the

[81] We made an exception if less than 24 subjects who completed the survey showed up to the experiment. In total, three subjects were allowed to participate despite not completing the online-survey.

[82] We follow the use of voice recordings to deliver instructions, as in Bartling, Engl and Weber (2015). This, combined with standardized instructions and computerized interfaces, ensures highly replicable environments across sessions.

total payoffs at the end of their decisions and never knew the precise payoffs for their choices.

Second, while their payment was determined and prepared, participants completed a survey in which they rated various firms and industries on a scale from 1 (*very immoral*) to 5 (*very moral*). For firms, clients also had the option to choose "I don't know this organization." The set of firms and industries were the same as the ones for which we elicited laboratory participants' willingness to work. The complete list of firms and industries is available in Appendix Tables D12 and D13.

5.5 Results

In presenting our results, we focus our attention on our first two hypotheses. We first study outcomes in the laboratory experiment, particularly outcomes in the labor market and their connection to the behavioral measure of concern for morality (θ^{Exp}). We then extend our measures to the survey data, to study the relationship between both of our measures of theta (θ^{Exp} and θ^{Sur}) and subjects' stated job market preferences. While choices in our experiment were conducted with respect to points, we present the results in terms of Swiss francs (CHF) to provide a clearer indication of the economic relevance.

5.5.1 Behavior in the laboratory

We first construct θ^{Exp} based on behavior in the behavioral task that subjects completed at the beginning of the laboratory session (Appendix Table D1 shows the distribution of choices in the behavioral task). Let m_{ir} be the number that individual i reports if the actual die roll is r. We classify an individual as *low-theta* if $m_{ir} \geq r$ for all $r \in \{1, 2, \ldots, 6\}$ and $m_{ir} > r$ for at least one r; that is, participant i is classified as having a low concern for morality if he or she lies at least once for personal gain and never in self-

harmful manner. We classify the remaining participants as *high-theta*.[83] Based on this classification, we have 66 (27.5 percent) low-theta types and 174 (72.5 percent) high-theta types.[84] Given our interpretation of θ, we will often refer to low-theta types as *immoral* and high-theta types as *moral types*. We next explore the differential behavior of these types in the labor markets and the consequences of this behavior.

Differential market behavior of moral and immoral types

Assuming that θ^{Exp} measures a stable concern for morality that translates into labor-market choices, then we should observe differential behavior in the laboratory labor markets between high-theta and low-theta types, but only when employment requires immoral work. This is confirmed in the data. In the immoral treatment, low-theta workers opted to submit wage requests more frequently than high-theta workers (90.6 percent vs. 61.6 percent). Moreover, 21.5 percent of the high-theta workers never participated in the market (i.e., refused to submit a wage request in any of the 15 periods), but this is true of only 4.3 percent of low-theta workers. By declining to submit a wage request, a subject indicates an unwillingness to do the work even at a wage of up to 50 CHF (1000 points), the highest possible wage request in our experiment. Conditional on submitting a wage request, low-theta types requested lower compensations (on average, CHF 3.56 vs. CHF 4.14).[85] In

[83] A total of 13 subjects (5.4 percent) harmed themselves with a lie ($m_{ir} < r$, e.g., reporting $m_{ir} = 1$ when $r = 2$). Since these subjects do not appear to be motivated by egoism, we classify them as high-theta. The remaining 161 subjects classified as high-theta always report the true number. Classifying subjects that lied in a self-harmful manner as high-theta types is conservative in that they act less morally than the honest subjects (see Table D3 in the Appendix). Results do not change if we drop these subjects or if we classify them instead as low-theta types.

[84] In principle, we could classify subjects into more than two categories, e.g. conditional on the number of lies. Due to the low number of subjects with different lying-patterns (see Appendix Table D1), we opt for a binary classification. Moreover, as Appendix Table D3 indicates, employment in the immoral treatment is generally higher for subjects who tell more lies.

[85] If we first calculate the average wage request by each individual (excluding observations in which the subject did not participate in the market), and then take the average of these averages, we obtain a mean wage request of CHF 4.36 for high-theta types and CHF 3.81 for low-theta types (t=-1.68, p-value = 0.104).

Appendix Table D2, we report the results of a two-stage hurdle regression of the decision of whether to submit a wage request and, conditionally, the actual wage request. The key independent variable is a subject's type from the behavioral task at the beginning of the experiment. In the immoral treatment, the model reveals a strong effect of a participant's type on the decision to participate (p-value<0.001). Furthermore, according to the model, low-theta types submit conditional reservation wage requests that are approximately 0.50 CHF lower (p<0.1). Moreover, these effects do not become weaker over time—if anything the coefficients indicate that the greater willingness of low-theta types to participate in the immoral labor market becomes stronger over time, though the magnitudes are small and statistically insignificant.[86] Hence, a subject's moral type seems to predict their willingness to seek employment in an immoral job.

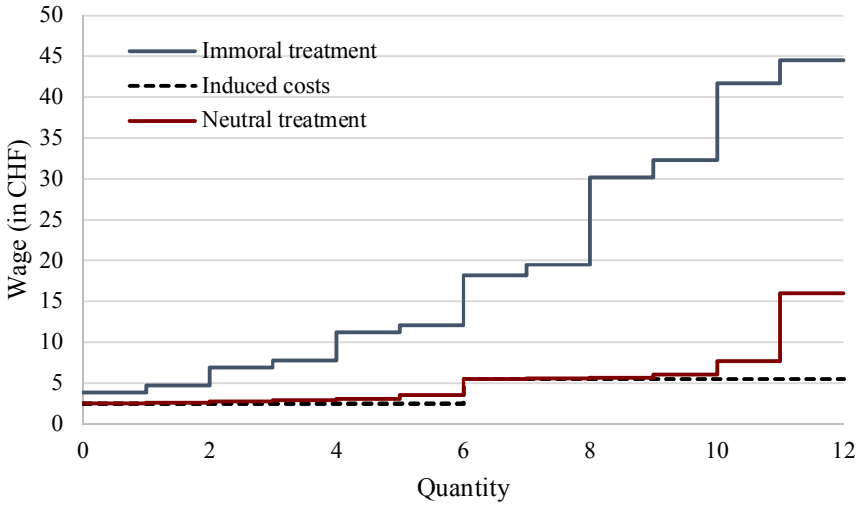
However, in the neutral treatment, non-participation is virtually non-existent—there were only 3 cases in total in which participants chose not to participate, representing 0.28 percent of all observations. That is, there is nearly universal participation when the job does not involve immoral behavior. Moreover, the wage requests of high-theta (CHF 2.91) and low-theta (CHF 2.89) types do not differ in magnitude or in statistical significance (p-value from hurdle model = 0.827).

Thus, there seems to be differential participation in the market between moral and immoral types when working involves immoral acts. In particular, high-theta types withdraw their participation and make higher wage requests when the market is immoral. However, when the work activity is neutral both types participate highly frequently. As a direct consequence of these observations, labor supply differs substantially between the two

[86] Specifically, if we add a linear time trend to the hurdle model and its interaction with theta (see Appendix Table D2), we find that low-theta types become slightly more likely to participate over time and provide lower reservation wages, relative to high-theta types. However, both coefficients are small and statistically insignificant.

kinds of markets, as shown in Figure 16.[87] In the neutral treatment, labor supply is fairly close to the theoretical prediction. However, for any given wage, there is a substantially lower supply of labor in the immoral treatment.

Figure 16: Labor supply for neutral and immoral work in the laboratory



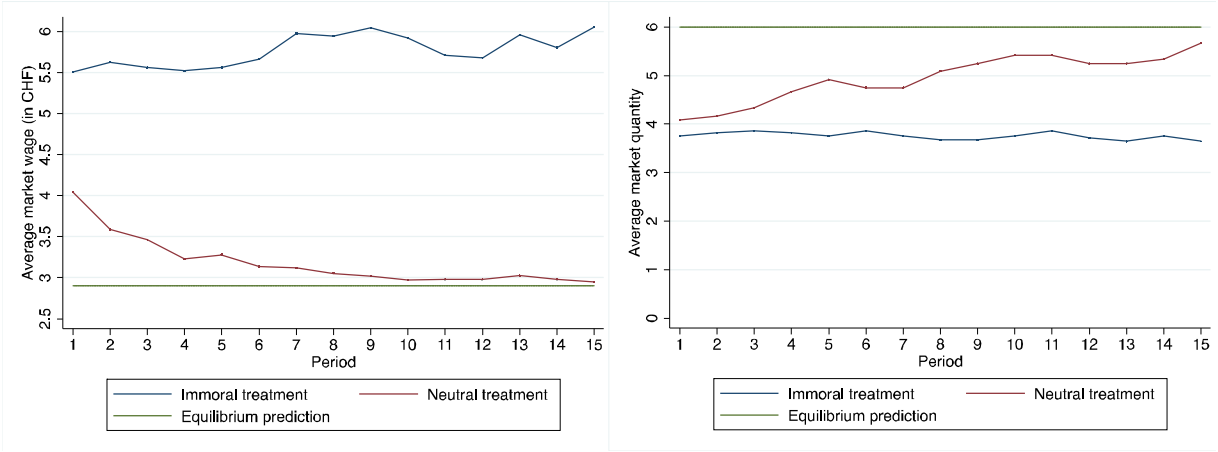
Outcomes in immoral labor markets

In this section, we explore the implications of the above heterogeneous behavior for labor market outcomes. As predicted by Proposition 1, we find a substantial immorality premium. This follows from the differential labor supply in Figure 16 and is shown in Figure 17a. While market wages in the neutral treatment converge toward the equilibrium prediction of CHF 2.90, the average market wage is persistently higher in the immoral treatment (t-test from a regression with standard errors clustered at the market-level, p-value<0.001). Hence, our laboratory labor market yields a substantial and persistent wage premium for immoral work in a setting in which only the morality of work varies. This

[87] Wage requests are censored at the maximal wage request that subjects could make (50 CHF). For this figure, we set the wage requests of subjects who are not willing to participate to CHF 50. Therefore, the supply curve for the immoral treatment should be interpreted as a lower bound. Figure D2 in the Appendix shows the labor supply if we only consider the last 5 periods.

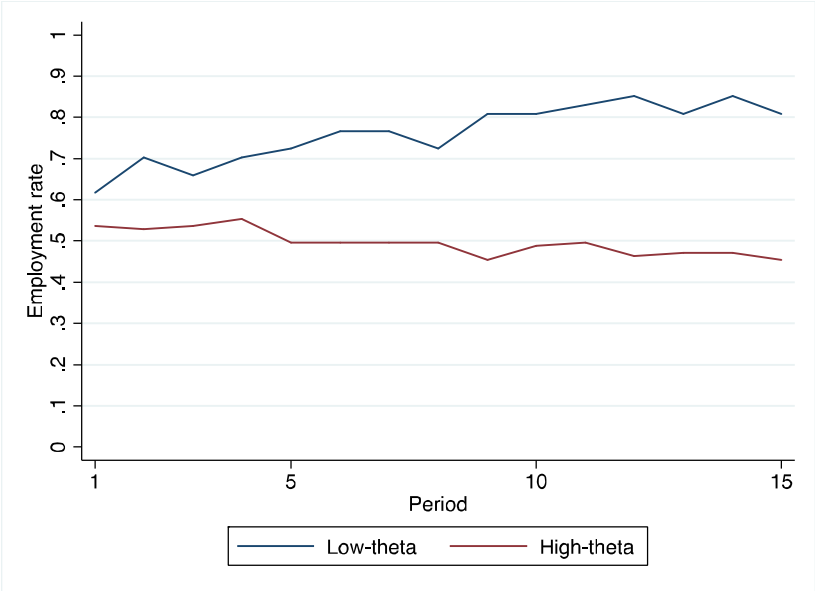
laboratory evidence corroborates the field evidence in support of Proposition 1 shown in Figure 13 and Table 15.

Figure 17: Immorality premium in laboratory labor markets



(a) Market wage (b) Market quantities

Figure 18: Employment rate by the two types in the immoral treatment



Of course, one concern might be that the immorality premium is the result of the specific labor demand structure that we employ in our design. However, as evident from the differences in labor supply that we identify in Figure 16, we would very likely obtain wage premiums under a wide variety of demand specifications.

We also find persistent differences in the employment levels in the two markets (Figure 17b). While the neutral market converges to the equilibrium prediction of 6, the average market quantity remains below 4 in the immoral markets. This difference is highly significant (t-test, p-value<0.001). Figure 17 provides further evidence that the manifestation of participants' morality does not erode over the course of the experiment.

We next turn to Proposition 2, which predicts that low-theta types will be hired in the immoral markets. Figure 18 shows that, indeed, high-theta types are consistently employed less frequently in the immoral treatment. Table 18 shows that, on average, low-theta types are 26.8 - 0.2 = 26.6 percentage points more likely to be employed than high-theta types (column 1). This difference is highly significant ($p < 0.001$) and robust to adding market fixed effects (column 2). Moreover, this observation is robust to other ways of constructing Θ^{Exp} from behavior in the behavioral task (see Table D3 in the Appendix). The results are similar if we use, as a dependent variable, the number of work units provided (0, 1 or 2) rather than a binary measure of employment (column 3 and 4). In the neutral treatment, we do not find a significant difference in employment rates between the two types (columns 1 to 4; see also Appendix Figure D3). This again suggests that the difference in hiring rates in the immoral treatment is driven by differences in concerns for morality and not some other difference between high- and low-theta types.

Propositions 3 and 4 predict heterogeneous treatment effects in terms of market income. To measure income, we use the sum of all earnings accumulated by a worker over

the 15 market periods. The propositions predict that these differences can be explained by the subjects' types and by the types of the other workers in their markets.

Table 18: Relationship between Θ^{Exp} and outcomes in the experimental labor markets

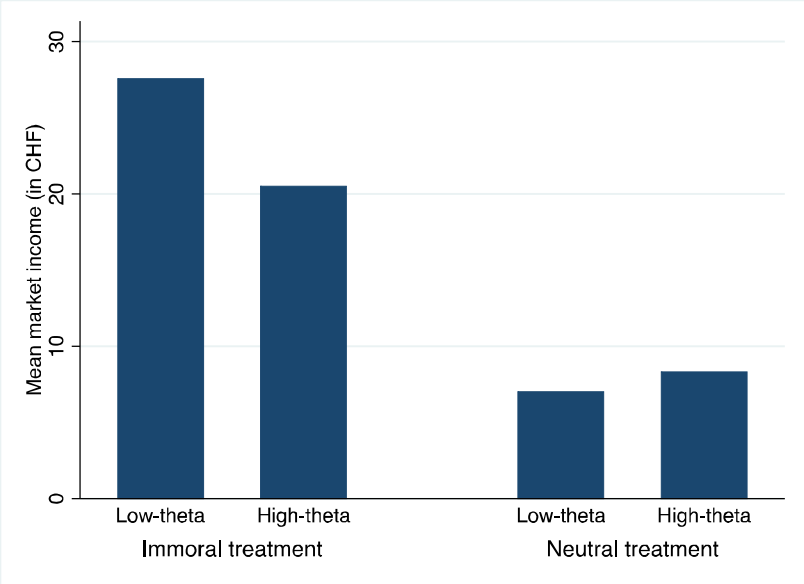
Dependent variable:	**Employment rate**		**Number of jobs**		**Market income**	
	(1)	(2)	(3)	(4)	(5)	(6)
Low-theta (θ_L^{Exp})	-0.002	-0.034	-0.002	-0.034	-1.33**	-0.56
	(-0.05)	(-0.90)	(-0.05)	(-0.90)	(-2.40)	(-1.39)
Immoral market	-0.331***		-0.272***		12.15***	
	(-7.70)		(-7.11)		(9.96)	
θ_L^{Exp} * Immoral market	0.268***	0.249***	0.256***	0.264***	8.42***	11.37***
	(4.54)	(3.58)	(4.10)	(3.57)	(2.71)	(3.64)
N	240	240	240	240	240	240
R^2	0.179	0.319	0.103	0.174	0.138	0.243
p-value: $\theta_L^{Exp} + \theta_L^{Exp}$*Im = 0	0.0000	0.0007	0.0000	0.0009	0.026	0.001
Market FE	No	Yes	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Independent variables: Low-theta in {0, 1}, Immoral market in {0, 1}. Standard errors clustered at market level; t-statistics in parentheses; * - p < 0.1; ** - p < 0.05; *** - p < 0.01.*

Specifically, Proposition 3 predicts that the immoral types benefit from an increase in the job's immorality. Figure 19 and results from linear regressions (see Table 18, Columns 5 and 6) show that the low-theta types earn more market income than the high-theta types in the immoral treatment, but not in the neutral treatment. The difference of CHF 7.09 (column 5) is statistically significant at the 5 percent level (p-value = 0.026). Note that the potential market income is constrained by the market wage. If we control for the market wage by adding market fixed effects (column 6), the immoral types are estimated to earn CHF 10.81 more than the moral types (p-value = 0.001). However, we find a different pattern in the neutral market: low-theta types earn slightly less (CHF 1.33) than the high-

theta types (p-value = 0.021) and this difference is not significant if we control for the market wage.

Figure 19: Market income by moral type

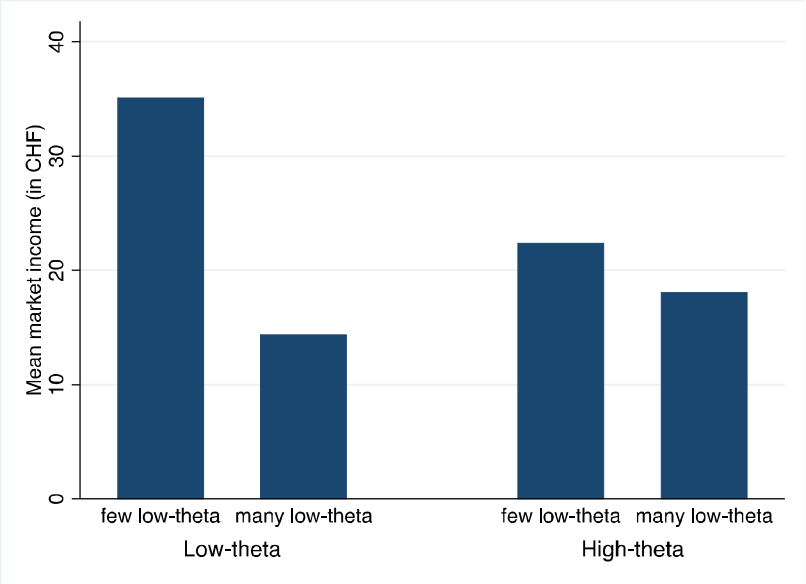


According to our theoretical analysis, doing an immoral job comes with a psychological cost—for example, from guilt or shame. Proposition 3 predicts that the market wage exceeds these psychological costs for the immoral types, resulting in a net benefit from an increase in immorality. Our design allows us to look at this type of welfare: in the experimental market, subjects submit their reservation wages.[88] The difference between the induced cost and the reservation wage is therefore an individual measure of the psychological cost for doing the immoral job in a specific period. For every job for which a subject is hired, we subtract her psychological cost for doing this job from her earnings. This provides a measure of the aggregated welfare, or utility, from market participation. We find a similar pattern as for income: In the immoral treatment, the low-theta types have a CHF 7.05 higher

[88] Note that this is an advantage of the uniform-price sealed-offer auction over the double auction. We also have a second welfare measure: subjects reported their happiness after the final market period. We do not find a statistically significant difference in happiness between types in both treatments (see Appendix Table D4).

corrected market income than the high-theta types (p-value = 0.007). In the neutral treatment, the difference between types is very small, and in the opposite direction (see Appendix Table D4).

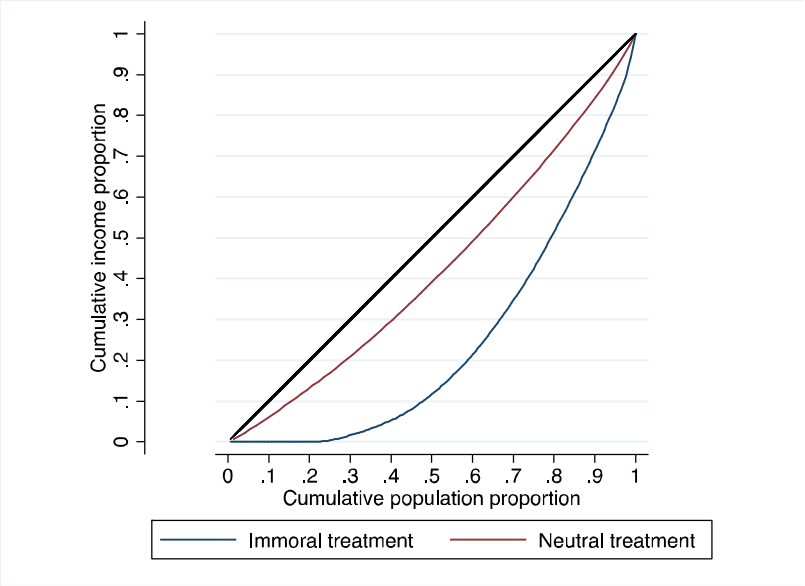
Figure 20: Externalities of moral behavior for immoral types



Finally, Proposition 4 predicts that, in the immoral treatment, immoral types have higher utility in the presence of more moral types. To test this prediction, for each subject we count the number of other workers of low-theta type in her market. We then split the sample based on the median of this measure: we classify a subject as being in a *few low-theta types market* if the number of (other) low-theta type workers is lower than 2, and as being in a *many low-theta types market* if the number of (other) low-theta type workers is 2 or more. This results in 70 subjects in the first category, and 98 subjects in the second category. The mean earnings of subjects in the immoral treatment, based on their own type and the median type of others in their market is shown in Figure 20. The income of high-theta types is CHF 4.33 higher in a market with few low-theta types than in a many low-theta types market (p-value = 0.051). However, for low-theta types, being in a few low-theta types market increases the income by an additional CHF 16.35 (p-value = 0.002),

resulting in a total difference of CHF 20.68 (p-value<0.001).[89] In Appendix Table D5 we confirm these observations in regression analysis or using our measure of welfare.

Figure 21: Income inequalities across treatments



Our findings confirm all four predictions from our simple model of heterogeneity in the morality of work and concerns for morality. We find evidence that people with high concern for morality consistently refuse to do immoral jobs (or require a high wage), thereby decreasing labor supply and increasing the equilibrium wage to produce an immorality premium. As a consequence, subjects with a low concern for morality are better off in immoral markets, in particular if they are in a market with many moral subjects.

One additional consequence of heterogeneous concerns for morality is that the income distribution differs substantially between the two treatments. While in the neutral treatment, income is almost equally distributed (Gini coefficient=0.15), we find substantial

[89] If we use the number of (other) low-theta types in the market instead of doing a median split, we find similar results. For high-theta types, the income is estimated to increase by CHF 2.45 per additional high-theta type in the market (p-value = 0.045). For low-theta types, the income increases by CHF 8.21 for every additional high-theta type (p-value = 0.014).

income inequality in the immoral treatment (Gini coefficient=0.38). Figure 21 shows the Lorenz curves for both treatments.

5.5.2 Survey evidence on moral concerns and labor market preferences

Our laboratory experiment provides support for all four of our predictions. This evidence indicates that heterogeneity in moral concerns, in combination with varying degrees of immorality in the nature of work, can produce stable wage premiums. Nevertheless, the extent to which heterogeneous concerns for morality affect labor-market outcomes outside the laboratory requires further evidence. Some of this evidence is evident in our analysis in section 5.5.2—and in related prior work by others—that provides support for Proposition 1. However, to our knowledge there remains a gap in showing the kind of sorting process underlying Proposition 2, that less moral people are more likely to go into work in immoral jobs. In this section, we attempt to reduce this gap using survey evidence.

First, we use the responses provided by our laboratory participants to the on-line survey that they completed roughly one week before the laboratory sessions.[90] This allows us to construct a second individual measure of concern for morality, Θ^{Sur}. Unlike the measure that is based on laboratory behavior (Θ^{Exp}), Θ^{Sur} is not tailored to measure θ in the specific context of our experiment. Moreover, as it is based solely on survey questions, it is easier to collect for future research and perhaps also represents a more general notion of morality than the behavioral task in our laboratory experiment. Moreover, a comparison of Θ^{Exp} and Θ^{Sur} provides some evidence on the stability of moral concerns across time and contexts, which is necessary for heterogeneous moral concerns to persistently influence labor market outcomes.

[90] Three subjects did not complete the online-survey. We exclude these subjects from the analysis.

Constructing Θ^{Sur}

Table 19 lists 9 subscales from the morality measures we collected as part of our survey. The definition and summary statistics for each of these is provided in the Appendix Table D6. Column 1 in Table 19 shows the coefficients from independent simple regressions of a subject's type measured by Θ^{Exp} on that particular subscale. The dependent variable is binary, indicating that a subject is a high-theta type. The results reported show a significant positive correlation between Θ^{Exp} and all personality measures, except for *Work ethics 1* and *Work ethics 2*. These positive relationships suggest that Θ^{Exp} and our measures of psychological traits capture similar individual characteristics.

Table 19: Items comprising Θ^{Sur} their relationship to Θ^{Exp}

	Regression coefficient of Θ^{Exp} (1)	**Factor loadings (weights for Θ^{Sur}) (2)**
Protected value 1	0.540***	0.664
	(3.66)	
Protected value 2	0.352**	0.708
	(2.22)	
Work ethics 1	-0.039	0.213
	(-0.39)	
Work ethics 2	0.042	0.252
	(0.52)	
HEXACO sincerity	0.362**	0.482
	(2.53)	
HEXACO fairness	0.353***	0.611
	(2.61)	
HEXACO greed avoidance	0.225*	0.477
	(1.73)	
HEXACO modesty	0.236*	0.508
	(1.79)	
Charity attitude index	0.711***	0.545
	(3.11)	

*Notes: (1): Coefficient estimates of linear probability models. N = 237 for each regression (3 subjects did not complete the online-survey and are excluded). Dependent variable: being a high-theta type according to Θ^{Exp}. Independent variables: survey measures in [0,1], higher numbers indicate more morality. Robust standard errors; t-statistics in parentheses; * - p < 0.1; ** - p < 0.05; *** - p < 0.01. (2): Factor loadings of survey measures on Θ^{Sur}.*

Next, we aggregate these nine psychological measures by performing a principal-component factor analysis. We selected the factor with the highest eigenvalue (eigenvalue = 2.44) to represent our *psychological* measure of concern for morality, i.e., Θ^{Sur}. Column 2 in Table 19 presents the corresponding factor loadings. We normalized Θ^{Sur} such that it lies between 0 and 1; the resulting variable has a mean of 0.5 and a median of 0.50. Low values represent a low concern for morality.[91] The distribution of Θ^{Sur} is presented in Appendix Figure D4.

Does Θ^{Sur} predict behavior in the laboratory?

We first consider whether the survey-based measure of concern for morality (Θ^{Sur}) has any relationship to the one constructed from the behavioral task in the laboratory experiment (Θ^{Exp}). Consistent with the positive relationship of the individual items shown in Table 19, a regression of Θ^{Exp} on Θ^{Sur} shows positive and significant relationships ($\mathsf{coefficient} = 0.723$, $p < 0.001$), i.e., a person who is characterized by a low concern for morality according to our psychological measures is more likely to lie self-servingly in the behavioral measure in the experiment. With a series of simple psychological questions, we are therefore able to partially predict people's tendency to act immorally in an incentivized context.

We next consider the extent to which Θ^{Sur} also predicts participants' behavior in immoral labor markets. Results from a linear regression of the employment rate on Θ^{Sur} indicate that participants with very low concern for morality ($\Theta^{Sur} = 0$) are 43.9 percentage points more likely to be hired in immoral labor markets than participants with a very high concern for morality ($\Theta^{Sur} = 1$). However, this difference is only marginally significant (p-value = 0.057, see Appendix Table D7, column 1). A less noisy measure of subjects'

[91] In the appendix, we demonstrate that our results robust to different aggregation mechanisms. Specifically, we look at two alternative aggregation mechanisms: i) each of the nine survey measures is given equal weight and ii) weight of the measures is determined by a regression of Θ^{Exp} on the survey measures.

behaviors in the market are their actual choices. Results from a hurdle model indicate that low-theta survey types ($\Theta^{Sur} = 0$) are 52.1 percentage points more likely to participate in immoral labor markets than individuals classified as very moral (p-value = 0.015, see Appendix Table D8). However, they do not have significantly lower reservation wages than the moral types in those markets (p-value = 0.548). In the neutral treatment, we do not find a significant difference in employment rates between the two types (see Appendix Table D7, columns 1 and 2).

Do Θ^{Sur} and Θ^{Exp}predict real-world labor market preferences?

Our study collects two measures of participants' concern for morality (Θ^{Sur} and Θ^{Exp}). Furthermore, the initial on-line survey elicits participants' willingness to work for several firms and industries. In addition, we obtained independent ratings of the perceived immorality of these firms and industries. In this section, we use all of this information to analyze how our measures of concern for morality connect to expectations about real labor market outcomes.

We create a measure of *perceived industry immorality* and of *perceived firm immorality* by averaging the ratings from subjects who also served in the role of clients and scaling them such that they lie between -1 (very moral) and +1 (very immoral), where 0 means neutral.[92] These variables are used as noisy measures of the immorality of work, $I(j)$, a key component of our theoretical model, where j represents an industry or a firm. The horizontal axis of Figures 22a and 22b plot the resulting normalized ratings for industries in our sample, the horizontal axis of Figures 22c and 22d plot the normalized ratings for firms. In Appendix

[92] Regarding *firms*, clients also had the option to choose "I don't know this organization" instead of rating the firm. To calculate the perceived firm immorality, we exclude these observations. Alternatively, we could code these as neutral ratings. These two measures are highly correlated (corr=0.9854). Our results do not change substantially if we use the alternative measure (see Table D10).

Tables D12 and D13, we provide a list of all the industries and firms listed in the survey, and the average perceived immorality.

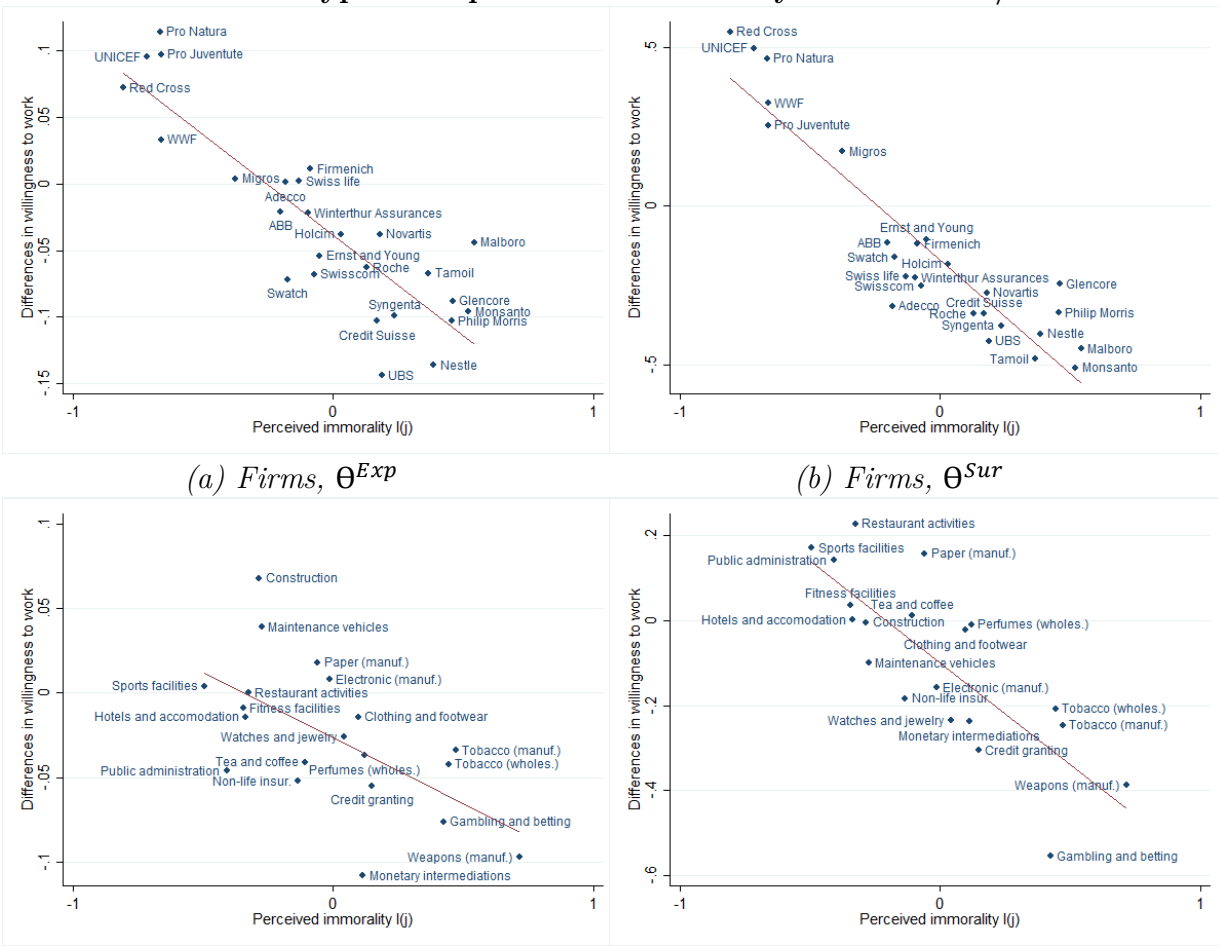
These ratings on the perceived immorality of industries and firms allow us to study how our measures of concerns for morality (Θ^{Sur} and Θ^{Exp}) correlate with the expected future labor market outcomes of subjects with varying degrees of concern for morality, measured by Θ^{Sur} and Θ^{Exp}. We normalized participants' stated willingness to work for firms and industries, such that they take values between 0 (*not at all willing*) and 1 (*very much willing*).[93]

The vertical axes of Figures 22a and 22b plot the difference in willingness to work for the industries between participants who were classified as moral or immoral, according to Θ^{Exp} (Figure 22a) or Θ^{Sur} (Figure 22b). The strong negative relationship indicates that participants classified as immoral are, on average, more willing to work for industries perceived as immoral. This tendency is robust to controlling for participants' gender, age, Swiss nationality and subject of study (see Appendix Figure D5).

More generally, Table 20, columns (1) to (4), shows that while there is little evidence for a systematic difference in willingness to work for neutral industries between moral and immoral types, participants classified as immoral are significantly more willing to work in industries that are classified as immoral. This pattern is significant at least at the 5%-level and holds for both of our measures of individual moral concerns, Θ^{Sur} and Θ^{Exp}.

[93] We miss information on willingness to work for the three participants who did not fill out the online survey. For the first session, we also miss willingness to work in five industries for five subjects that filled out the survey. We exclude all these missing observations from the analysis. Regarding willingness to work for *firms*, participants also had the option to choose "I don't know this organization." This option was chosen in 17.8 percent of all answers. We also exclude these observations. We obtain similar results if we classify such observations as "indifferent" or if we restrict our analysis to subjects that know all firms (N=52), see Table D11.

Figure 22: Correlation between the difference in willingness to work between moral and immoral types and perceived immorality of industries/firms



(a) Firms, Θ^{Exp} *(b) Firms,* Θ^{Sur}

(c) Industries, Θ^{Exp} *(d) Industries,* Θ^{Sur}

Source: Survey study (Perceived immorality), on-line survey (Willingness to work, Θ^{Sur}*), Laboratory experiment (*Θ^{Exp}*). Notes: Differences in willingness to work: Coefficient estimates of linear regression models of the participants' willingness to work for different industries (a and b) or firms (b and c) on* Θ_H^{Exp} *(a and c) or* Θ_H^{Sur} *(b and d). Dependent variable: Willingness to work is in {0, 0.25, 0.5, 0.75, 1} where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm ("I don't know this organization") or did not fill out the questionnaire are excluded. Independent variables: a and c use* Θ^{Exp} *to classify participants, where* Θ^{Exp} *=0 for low- theta types and* Θ^{Exp} *=1 for high-theta types, while b and d use* Θ^{Sur} *in [0,1] instead. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Standard errors clustered at individual level.*

Table 20: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types

Dependent variable:	**Willingness to work for industry *j***				**Willingness to work for firm *j***			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Perceived immorality (I(j))	-0.232***	-0.232***	-0.050	-0.050	-0.140***	-0.139***	0.114**	0.110**
	(-7.05)	(-7.04)	(-1.01)	(-1.00)	(-4.30)	(-4.32)	(2.51)	(2.43)
Type experiment (Θ^{Exp})	-0.027	-0.028			-0.043	-0.0508*		
	(-1.28)	(-1.40)			(-1.54)	(-1.91)		
$\Theta^{Exp} * I(j)$	-0.078**	-0.078**			-0.154***	-0.154***		
	(-2.11)	(-2.10)			(-4.06)	(-4.11)		
Type survey (Θ^{Sur})			-0.101*	-0.107*			-0.173**	-0.211***
			(-1.74)	(-1.72)			(-2.23)	(-2.78)
$\Theta^{Sur} * I(j)$			-0.479***	-0.479***			-0.731***	-0.722***
			(-5.12)	(-5.11)			(-8.37)	(-8.27)
N	4715	4715	4715	4715	5064	5064	5064	5064
R^2	0.098	0.119	0.105	0.127	0.110	0.141	0.128	0.160
Control variables	No	Yes	No	Yes	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Dependent variable: Willingness to work is in {0, 0.25, 0.5, 0.75, 1} where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm ("I don't know this organization") or did not fill out the questionnaire are excluded. Independent variables: (1), (2), (5) and (6) use Θ^{Exp} to classify participants, where Θ^{Exp}=0 for low- theta types and Θ^{Exp}=1 for high-theta types, while (3), (4), (7) and (8) use Θ^{Sur} (in [0,1]) instead. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study. Standard errors clustered at individual level; t-statistics in parentheses; * p < 0.1; ** p < 0.05; *** p < 0.01.*

We repeat the same analysis using data on individuals' willingness to work for our selection of well-known *firms* in Switzerland. The vertical axis of Figures 22c and 22d plot the difference in willingness to work for the firms between participants who were classified as moral and immoral according to Θ^{Exp} (Figure 22c) or Θ^{Sur} (Figure 22d). Again, participants that are classified as immoral types are on average more willing to work for firms perceived as immoral. This relationship is confirmed by Table 20, columns (5) to (8): immoral participants are more willing to work for firms that other people rate as more immoral ($p < 0.01$), which is again true for both measures of concern for morality (Θ^{Sur} and Θ^{Exp}). This finding indicates that firms that are perceived as immoral differentially attract applicants with a lower concern for morality. As we show in Appendix, the

differential willingness to work for immoral firms and industries of moral and immoral types does not depend on how we construct θ^{Sur} (Table D9), and are also consistent to using an ordered logit specification that better accounts for the nature of our dependent variable (Table D14).

The above analysis provides evidence consistent with Proposition 2 in real labor markets. Those who are least concerned with morality—low-theta types—are significantly more willing to work in firms that are perceived as less moral. Moreover, this provides us with a revealed (or stated) preference method for classifying the immorality of firms or industires: by immoral individuals' willingness to be employed there.

Finally, in the on-line survey, we also asked participants to rate how much they expect to earn when they reach the age of 40. We do not find any statistically significant correlation between participants' type and their earnings' expectations, although a regression of expected future wages on θ^{Sur} reveals a positive relationship such that the lowest types ($\theta^{Sur} = 0$) report expected income that is 30,272 CHF higher on average than the expected income of the most moral type ($\theta^{Sur} = 1$, p = 0.125). This relationship is consistent with Proposition 1. However, the earnings expectations measures for such a long time horizon—on average 18 years—are perhaps less reliable than the more contemporaneous statements of willingness to work for different firms.

5.6 Conclusion

We investigate how individual heterogeneity in concerns for morality affects outcomes in "immoral" labor markets. Our study employs a combination of laboratory and survey measures to identify heterogeneity in concerns for morality and to measure or create variation in the immorality of jobs. We use these different kinds of data to test two main hypotheses—first, that an immoral labor market will yield immorality wage premiums (and

reduced quantities) and, second, that immoral types will be more likely to be hired in immoral labor markets.

In a laboratory setting, we use a simple behavioral task to classify individuals into "moral" and "immoral" types (Θ^{Exp}). We then show that this characteristic predicts the outcomes that individuals obtain as we experimentally vary *only* the immorality ($I(j)$) of a labor market. We find confirmation of both our hypotheses. Immoral labor markets yield significantly higher wages and reduced quantities. Moreover, immoral workers are significantly more likely to be hired in an immoral labor market than moral workers, but this difference disappears in a neutral labor market. We also find that a market for immoral work overcompensates the immoral types who are hired, particularly when there are many moral types in their market.

We separately use survey responses to classify the immorality of real-world firms and industries ($I(j)$) and demonstrate that industries classified as immoral produce higher wages. We also use surveys to obtain a separate measure of workers' moral types (Θ^{Sur}). This individual characteristic is correlated with Θ^{Exp} and predicts subjects' behaviors in the laboratory labor market. Moreover, the survey-based measure also predicts labor-market preferences regarding working in jobs and industries that vary in their morality ($I(j)$). Workers who are less concerned with morality—in either the behavioral or survey-based measures—are more willing to work for firms that others regard as less moral.

Despite widespread intuition, we know of no other evidence that makes the above connections. Given the significance of many social ills produced by immoral work practices, our study sheds important new light on the interaction between individual's types, their willingness to do immoral work and the resulting labor-market outcomes. Our work has potentially important policy implications. For instance, in those jobs and industries with the greatest potential to do societal harm, we should want to see those people inclined to

act most socially responsibly. However, our evidence suggests that it is the *least* moral types who will sort into these industries.

Our findings also contribute to recent studies on how income relates to moral behavior. There is evidence that rich people behave less morally than others (Auten et al., 2002; Piff et al., 2010; Piff et al., 2012; Wang and Murnighan, 2014).[94] It remains unclear whether a rise in income makes people less moral, or if less moral people obtain higher incomes (see also Bartling et al., 2018). Our evidence is in line with the latter interpretation.

Our results also add to the literature on the erosion of morality in markets. On the aggregate level, erosion of morality in the context of experimental consumption markets with externalities has been investigated before (e.g. Falk and Szech, 2013; Bartling, Weber and Yao, 2015; Kirchler, Huber, Stefan and Sutter, 2016) with mixed findings. In our experiment, we find little evidence of erosion of morality over periods both on aggregate and at the individual level. Instead we observe that subjects persistently behave consistently with their moral type. This is particularly surprising, given that moral workers forgo very large sums of money, relative to the typical earnings in a laboratory experiment and directly observe others making relatively large amounts of money. This suggests that our measure of concern for morality captures a stable personality trait and the results indicate a stable concern of some individuals for moral behavior.

[94] However, there are some limitations in the existing evidence, see Andreoni, Nikiforakis and Stoop (2017).

Appendix A – Psychological Depreciation and Food Waste

Tables

Table A-1: Descriptive statistics of the relevant variables

Variable	N	Mean	SD	Min	Max	Description
Taste	259	5.35	1.52	2	9	Is 1 if evaluates the taste of the sandwich as extremely bad, and 9 if evaluates it as extremely good, in {1, 2, 3, … 9}
WTP	222	0.77	0.5	0	2	Willingness-to-pay for the second half of the sandwich, in [0; 2.2] CHF
Weight	259	12.68	24.47	0	78	Weight (in grams) of the unfinished first sandwich half, in [0; 82]
First half is entirely eaten	259	0.77	0.42	0	1	Is 1 if has entirely eaten the first half of the sandwich, and 0 otherwise
Eager	258	5.31	1.91	1	9	Is 1 if is not at all eager to eat the sandwich, and 9 if is extremely eager to eat it, in {1, 2, 3, … 9}
Appear	259	4.67	1.49	1	9	Is 1 if finds that the sandwich looks not at all appetizing, and 9 if finds it extremely appetizing, in {1, 2, 3, … 9}
Smell	259	5.93	1.33	2	9	Is 1 if finds that the sandwich smells really bad, and 9 if finds that it smells really good, , in {1, 2, 3, … 9}
Guess	128	4.12	2.05	1	9	Is 1 if thinks that the sandwich was definitely made 1 day ago, is 5 if can't really tell, and is 9 if thinks that it was definitely made 7 days ago, in {1, 2, 3, … 9}
Hungry	259	5.39	1.93	1	9	Is 1 if is not at all hungry, is 9 if extremely hungry, in {1, 2, 3, … 9}

See next page for the rest of the table.

Already eat.	259	0.63	0.48	0	1	Is 1 if has already eaten this type of sandwich in the past, 0 otherwise
Female	259	0.52	0.5	0	1	Is 1 if is a woman, 0 otherwise
Age	259	22.49	3.09	18	35	Age
Food pois.	259	0.28	0.45	0	1	Is 1 if has already experienced a serious case of food poisoning in the past, 0 otherwise
Swiss	259	0.79	0.41	0	1	Is 1 if the person has the Swiss nationality, 0 if the person is a foreigner
Educ	259	1.44	0.7	1	4	Highest level of education achieved (1: *Matura/Secondary*, 2: *Bachelor*, 3: *Master*, 4: *PhD*)
Heard	259	0.03	0.18	0	1	Is 1 if had heard of the study before participating, 0 otherwise

Table A-2: Tests of the difference in means and distributions of appearance and smell ratings between old and fresh sandwiches in the *Blind* conditions

	Appearance		**Smell**	
	Old	**Fresh**	**Old**	**Fresh**
N	63	65	63	65
Mean	4.714	4.877	5.857	5.831
Difference: Fresh -	0.163		-0.026	
p-values:				
Two-sided t-test	0.518		0.911	
Two-sided rank-sum test	0.448		0.988	

Notes: Only observations from the Blind conditions.

Table A-3: Relationship between the categorized taste ratings and the treatments

Dependent variable:	**Taste cat.**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Old	0.0237	0.0570	0.0628	0.0425	0.0856	0.102
	(0.15)	(0.37)	(0.41)	(0.20)	(0.40)	(0.48)
Salient	-0.0923	-0.0880	-0.0950	-0.121	-0.122	-0.140
	(-0.60)	(-0.57)	(-0.61)	(-0.60)	(-0.59)	(-0.68)
Salient*Old	-0.0412	-0.0959	-0.0855	-0.0627	-0.136	-0.137
	(-0.19)	(-0.43)	(-0.38)	(-0.21)	(-0.46)	(-0.45)
Already eat.		0.203*	0.189		0.276*	0.257*
		(1.77)	(1.62)		(1.81)	(1.66)
Hungry		0.0468	0.0570*		0.0634	0.0785*
		(1.51)	(1.81)		(1.51)	(1.83)
Food pois.		-0.0318	-0.00476		-0.0404	-0.00487
		(-0.27)	(-0.04)		(-0.26)	(-0.03)
Female		-0.0617	-0.0643		-0.0764	-0.0768
		(-0.56)	(-0.58)		(-0.52)	(-0.51)
Age		-0.0320*	-0.0270		-0.0417*	-0.0341
		(-1.79)	(-1.44)		(-1.80)	(-1.39)
Constant	2.246***	2.622***	2.574***			
	(20.90)	(5.76)	(5.21)			
N	259	259	259	259	259	259
R^2	0.00418	0.0394	0.0664	0.00214	0.0196	0.0336
Session FE	No	No	Yes	No	No	Yes

*Notes: Columns 1 to 3: Coefficient estimates of linear regression models. Columns 4 to 6: Estimated marginal effects of ordered Probit models. Dependent variable: Taste cat.=1 if taste ratings<5, =2 if taste ratings=5, =3 if taste ratings>5. Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. Estimated cuts for ordered Probit models are omitted from the table, McFadden pseudo R^2 are provided in columns 4 to 6, robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table A-4: Relationship between the willingness-to-pay for the second sandwich half and the treatments

Dependent variable:	Willingness-to-pay for the second half					
	(1)	(2)	(3)	(4)	(5)	(6)
Old	-0.128	-0.114	-0.105	-0.159	-0.141	-0.132
	(-1.41)	(-1.26)	(-1.16)	(-1.52)	(-1.37)	(-1.31)
Salient	-0.130	-0.143	-0.135	-0.182*	-0.198*	-0.189*
	(-1.44)	(-1.57)	(-1.53)	(-1.68)	(-1.83)	(-1.84)
Salient*Old	0.143	0.119	0.124	0.187	0.160	0.165
	(1.07)	(0.90)	(0.95)	(1.17)	(1.02)	(1.08)
Already eat.		0.0793	0.0852		0.0771	0.0854
		(1.17)	(1.30)		(0.96)	(1.12)
Hungry		0.0419**	0.0437**		0.0517**	0.0528**
		(2.36)	(2.41)		(2.33)	(2.38)
Food pois.		0.0161	0.0444		0.00571	0.0407
		(0.22)	(0.61)		(0.06)	(0.47)
Female		-0.0637	-0.0646		-0.0710	-0.0761
		(-0.96)	(-0.94)		(-0.89)	(-0.95)
Age		-0.00299	-0.00936		-0.00298	-0.0103
		(-0.26)	(-0.81)		(-0.22)	(-0.77)
Constant	0.863***	0.679**	0.622**	0.844***	0.614*	0.515
	(14.02)	(2.42)	(2.01)	(12.48)	(1.78)	(1.34)
Sigma				0.583***	0.572***	0.553***
				(17.13)	(16.95)	(16.62)
N	222	222	222	222	222	222
R^2	0.0118	0.0451	0.108	0.00792	0.0257	0.0621
Session FE	No	No	Yes	No	No	Yes

*Notes: Columns 1 to 3: Coefficient estimates of linear regression models. Columns 4 to 6: Estimated marginal effects of Tobit models. Dependent variable: WTP in [0; 2.2] CHF (the 37 observations with a negative WTP are excluded from the analysis). Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. McFadden pseudo R^2 are provided in columns 4 to 6, robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table A-5: Relationship between the categorized willingness-to-pay and the treatments

Dependent variable:	WTP cat.					
	(1)	(2)	(3)	(4)	(5)	(6)
Old	0.0508	0.106	0.0914	0.0460	0.165	0.136
	(0.40)	(0.88)	(0.76)	(0.19)	(0.68)	(0.55)
Salient	-0.0615	-0.0677	-0.0750	-0.167	-0.164	-0.185
	(-0.48)	(-0.54)	(-0.61)	(-0.72)	(-0.69)	(-0.77)
Salient*Old	-0.120	-0.191	-0.176	-0.137	-0.310	-0.294
	(-0.66)	(-1.10)	(-1.01)	(-0.43)	(-0.97)	(-0.89)
Already eat.		0.145	0.174*		0.249	0.294*
		(1.54)	(1.88)		(1.47)	(1.73)
Hungry		0.0839***	0.0772***		0.149***	0.141***
		(3.46)	(3.15)		(3.58)	(3.34)
Food pois.		-0.0306	0.0148		-0.0384	0.0499
		(-0.33)	(0.15)		(-0.23)	(0.29)
Female		-0.164*	-0.200**		-0.296*	-0.385**
		(-1.91)	(-2.27)		(-1.86)	(-2.37)
Age		-0.0151	-0.0151		-0.0261	-0.0319
		(-1.03)	(-1.00)		(-1.06)	(-1.21)
Constant	2.600***	2.483***	2.263***			
	(27.35)	(6.76)	(5.48)			
N	259	259	259	259	259	259
R^2	0.00863	0.0853	0.121	0.00580	0.0524	0.0777
Session FE	No	No	Yes	No	No	Yes

*Notes: Columns 1 to 3: Coefficient estimates of linear regression models. Columns 4 to 6: Estimated marginal effects of ordered Probit models. Dependent variable: WTP cat.=1 if WTP<0, =2 if WTP=0, =3 if WTP>0. Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. Estimated cuts for ordered Probit models are omitted from the table, McFadden pseudo R^2 are provided in columns 4 to 6, robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table A-6: Relationship between the decision to finish the first half of the sandwich entirely and the treatments

Dependent variable:	First half entirely eaten		
	(1)	(2)	(3)
Old	0.118	0.164**	0.163**
	(1.58)	(2.39)	(2.37)
Salient	0.0769	0.0733	0.0732
	(1.00)	(1.05)	(1.05)
Salient*Old	-0.145	-0.200**	-0.203**
	(-1.38)	(-2.09)	(-2.11)
Already eat.		0.149***	0.158***
		(2.91)	(3.10)
Hungry		0.0632***	0.0613***
		(4.52)	(4.30)
Food pois.		-0.0109	-0.00350
		(-0.20)	(-0.06)
Female		-0.233***	-0.246***
		(-4.89)	(-4.94)
Age		-0.00714	-0.00754
		(-0.82)	(-0.81)
Constant	0.708***	0.551**	0.456*
	(12.45)	(2.50)	(1.73)
N	259	259	259
R^2	0.0102	0.205	0.221
Session FE	No	No	Yes

*Notes: Coefficient estimates of linear probability models. Dependent variable: First half is entirely eaten in {0;1}. Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table A-7: Relationship between individuals' eagerness to eat the sandwich and the treatments

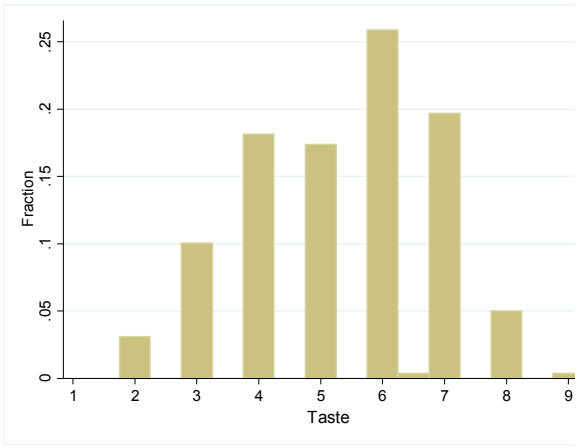
Dependent variable:	Eager					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Old	-0.0397	0.000542	-0.0430	-0.0246	-0.00466	-0.0319
	(-0.13)	(0.00)	(-0.15)	(-0.14)	(-0.03)	(-0.18)
Salient	0.303	0.341	0.351	0.210	0.235	0.250
	(0.99)	(1.11)	(1.19)	(1.14)	(1.24)	(1.35)
Salient*Old	-0.609	-0.694*	-0.639	-0.405	-0.464*	-0.447*
	(-1.44)	(-1.66)	(-1.54)	(-1.63)	(-1.85)	(-1.77)
Hungry	0.438***	0.431***	0.413***	0.259***	0.262***	0.258***
	(7.71)	(7.73)	(7.23)	(6.87)	(6.95)	(6.64)
Already eat.		0.666***	0.731***		0.406***	0.459***
		(2.98)	(3.40)		(2.95)	(3.34)
Food pois.		-0.381	-0.335		-0.219	-0.194
		(-1.64)	(-1.38)		(-1.55)	(-1.29)
Female		-0.314	-0.327		-0.175	-0.191
		(-1.47)	(-1.51)		(-1.34)	(-1.42)
Age		-0.00634	-0.0113		-0.00430	-0.00792
		(-0.18)	(-0.32)		(-0.20)	(-0.35)
Constant	2.965***	2.980***	3.212***			
	(7.91)	(3.25)	(3.16)			
N	258	258	258	258	258	258
R^2	0.206	0.248	0.300	0.0571	0.0696	0.0859
Session FE	No	No	Yes	No	No	Yes

*Notes: Columns 1 to 3: Coefficient estimates of linear regression models. Columns 4 to 6: Estimated marginal effects of ordered Probit models. Dependent variable: Eager=1: Not at all eager to eat the sandwich, =9: Extremely eager to eat the sandwich. Independent variables: Old, Salient, and Female are in {0, 1}; Already eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Food pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Hungry=1: Not at all hungry, =9: Extremely hungry. Estimated cuts for ordered Probit models are omitted from the table, McFadden pseudo R^2 are provided in columns 4 to 6, robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

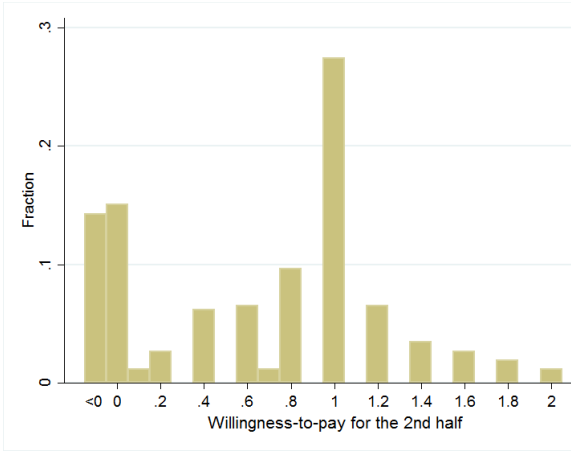
Figures

Figure A-1: Picture of the sandwich half distributed to participants

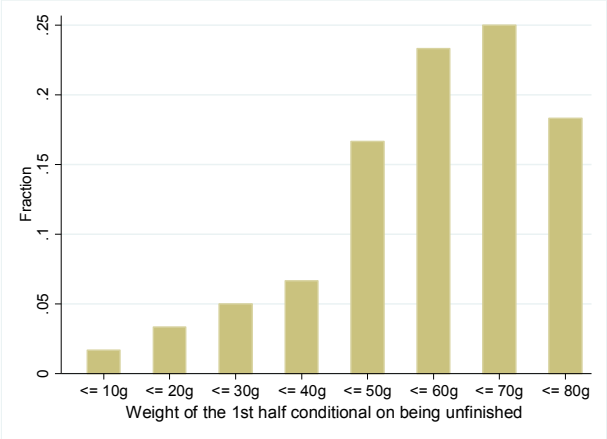
Figure A-2: Distribution of (a) taste ratings, of (b) willingness-to-pay for the second half and of (c) the weight of unfinished first halves



(a)



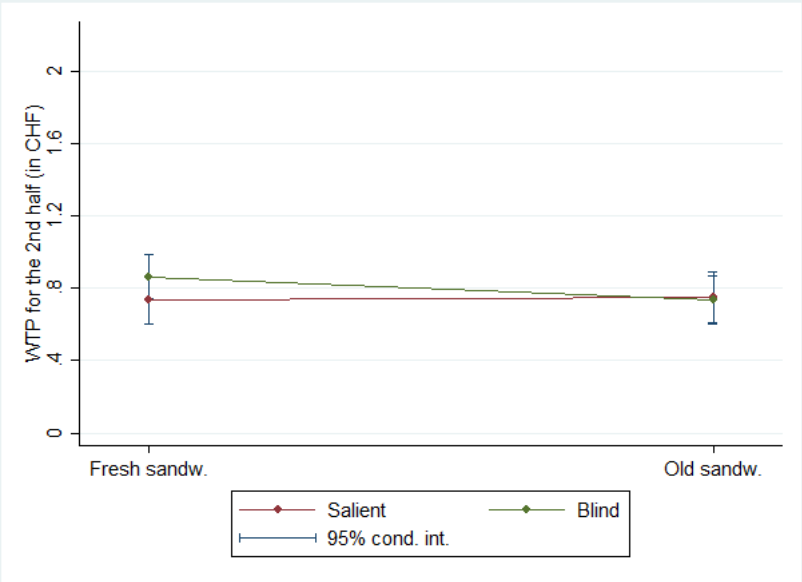
(b)



(c)

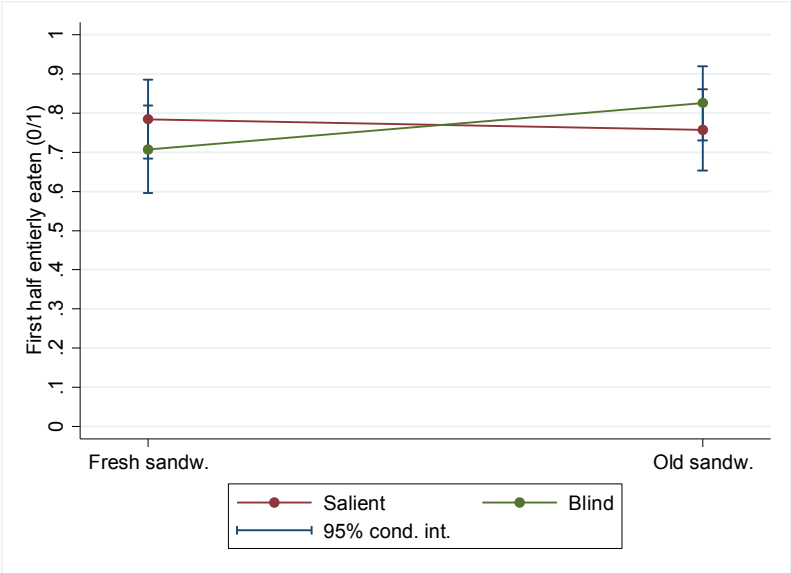
Notes: Taste=1: Extremely bad taste, =9: Extremely good taste; WTP in [-0.1; 2.2] CHF, where WTP=-0.1 if a person does not even want the second half for free. N=259 with N_{BO}*=65,* N_{BF}*=63,* N_{SO}*=65 and* N_{SF}*=66.*

Figure A-3: Average willingness-to-pay in each of the four conditions



Notes: WTP in [0; 2.2] CHF. N=222 (the 37 observations with a negative WTP have been excluded from the sample) with N_{BO}*=57,* N_{BF}*=54,* N_{SO}*=54 and* N_{SF}*=57.*

Figure A-4 : Probability to finish the first sandwich half per condition



Notes: N=259 with N_{BO}*=65,* N_{BF}*=63,* N_{SO}*=65 and* N_{SF}*=66.*

Instructions

Overview:

Instructions for the Salient-Old treatment are presented below. The Part's number (I, II or III) is indicated in the top-right corner of every page of instructions (e.g. Part I corresponds to "1/3"). Instructions for the Salient-*Fresh* condition differed only in the production date indicated in Part I. Instructions for the two *Blind* conditions did not display any information on the production date in Part I, and question 7 looked as follows:

In this study, half of the participants receive a sandwich that was made 1 day ago, and the other half of the participants receive a sandwich that was made 7 days ago.

7. How many days ago do you think that the sandwich that you received was made?

Definitely 1 day ago 1	2	3	4	I really can't tell 5	6	7	8	Definitely 7 days ago 9
O	O	O	O	O	O	O	O	O

Instructions for the Salient-Old condition:

ID: A1

Universität Zürich[UZH]

Instructions

You will receive one half of a pre-packaged sandwich purchased in a supermarket. We will open the package containing the sandwich right before giving it to you. The half of a sandwich that you will receive:

- was made on the 14th of March, **7 days ago**,
- weighs **80 grams**, and
- contains the following main **ingredients**: wheat, rye, barley, eggs, mayonnaise, mustard, and yoghurt.

Once you receive the sandwich, please do not start eating it yet. Take a few seconds to observe it and to answer the following questions by writing an "x" in the appropriate circle.

*1. How would you rate the **appearance of the sandwich** that you have just received?*

Not at all appetizing 1	2	3	4	5	6	7	8	Extremely appetizing 9
O	O	O	O	O	O	O	O	O

*2. How well does the sandwich **smell**?*

Smells very bad 1	2	3	4	Does not smell 5	6	7	8	Smells very good 9
O	O	O	O	O	O	O	O	O

*3. How **eager are you to eat the sandwich** that you received?*

Not at all eager 1	2	3	4	5	6	7	8	Extremely eager 9
O	O	O	O	O	O	O	O	O

To confirm that you read the text at the beginning of this page, please answer the following questions:

*Circle the ingredient that the sandwich **does not contain**:* *wheat* *eggs* *cucumber*

*Indicate the **weight** of the half of a sandwich that you received: …….. grams*

*Indicate **how many days ago** was your sandwich made: …….. day(s) ago*

You may now start eating your sandwich. **After you have taken at least 3 regular-sized bites**, please raise your hand to receive the next page of the questionnaire. At that point, you may continue eating as much of your sandwich as you want while completing the rest of the study.

ID: A1

Universität Zürich[UZH]

Questionnaire – Part I

*4. Please rate the **taste** of your sandwich.*

Extremely bad 1	2	3	4	5	6	7	8	Extremely good 9
O	O	O	O	O	O	O	O	O

5. As we told you earlier, you will receive 10 CHF for your participation in this study. With this money, you will also have the possibility to buy the 2nd half of the same sandwich (i.e., an additional 80 gr.) at a randomly-drawn price. For every possible price listed below, please indicate with an "x" whether or not you would like to buy the 2nd half of the same sandwich. At the end of the experiment, you will roll a die that will determine the actual price. Your answer for this actual price in the table below will determine whether or not you will pay this price out of your 10 CHF to receive the 2nd half of the sandwich.

If the die roll is…	…then the price of the 2nd half of the sandwich will be…	I want to buy the 2nd half of the sandwich at this price	I do not want to buy the 2nd half of the sandwich at this price
1	0.00 CHF	O	O
2	0.20 CHF	O	O
3	0.40 CHF	O	O
4	0.60 CHF	O	O
5	0.80 CHF	O	O
6	1.00 CHF	O	O
7	1.20 CHF	O	O
8	1.40 CHF	O	O
9	1.60 CHF	O	O
10	1.80 CHF	O	O
11	2.00 CHF	O	O
12	2.20 CHF	O	O

*6. Please rate the extent to which you consider your sandwich **healthy and nutritious** (vitamins, minerals, etc.)?*

Not at all healthy & nutritious 1	2	3	4	5	6	7	8	Extremely healthy & nutritious 9
O	O	O	O	O	O	O	O	O

Once you have answered all the questions of "Questionnaire – Part I," please raise your hand and wait for a staff member to provide you with the last part of the questionnaire.

ID: A1

Universität Zürich[UZH]

Questionnaire – Part II

3/3

In this study, half of the participants receive a sandwich that was made 1 day ago, and the other half of the participants receive a sandwich that was made 7 days ago. At the beginning of the instructions, we told you how many days ago the sandwich that you received was made.

*7. Do you remember **how many days ago** your sandwich was made? The sandwich that I received was made…*

1 day ago:	O
7 days ago:	O

*8. Have you eaten this specific type of pre-packaged sandwich **in the past**?* *Yes: O* *No: O*

*9. How frequently do you **eat sandwiches**?*

Never	Rarely (1 time per month or less)	Regularly (2 or 3 times per month)	Often (about 1 to 2 times per week)	Very often (more than 2 times per week)
O	O	O	O	O

*10. What **type(s) of sandwiches** do you usually eat? (You may select multiple answers.)*

O	Home-made
O	Freshly prepared (by a bakery, at a deli counter…)
O	Packaged sandwiches

11. When you eat a sandwich, how much do you care about how healthy and nutritious it is?

I do not care at all 1	2	3	4	5	6	7	8	I care a lot 9
O	O	O	O	O	O	O	O	O

12. How frequently do you prepare meals using raw ingredients? (Heating a frozen or prepackaged meal does not count.)

Never	Rarely (up to 3 times per month)	Regularly (about 1 to 2 times per week)	Often (more than 2 times per week)	Daily or almost daily
O	O	O	O	O

*13. Imagine that you want to eat a piece of hard cheese. You take it out of your refrigerator and notice that **there is some mold on it**. What do you do? (Please select the answer that you think would best represent your behavior.)*

I eat it.	O
I remove only the moldy part (i.e., as little as possible) to maximize the quantity of cheese left to eat.	O
I cut out any part that is at all close to the mold and eat the rest.	O

Universität Zürich[UZH]

I throw the piece of cheese away.	O

14. Below is a list of food items. Imagine that you bought these items fresh and stored them in an appropriate environment (e.g., in the refrigerator). Please indicate by an "x" in the corresponding cell (a) how safe you think it is to eat this item and (b) how likely you would be to eat it (assuming that you generally eat this kind of food even if you usually do not). Assume that each item smells and looks fine when you inspect or open it.

	How safe do you think it is to eat this item?			How likely would you be to eat this item?		
Food item description:	Not at all safe	Moderately safe	Completely safe	Not at all likely	Moderately likely	Very likely
A delivered (reheated) pizza with vegetables and meat that has been sitting in your fridge for 3 days						
An unopened plain (i.e., unflavored) yoghurt that expired 10 days ago						
Hard boiled eggs that you cooked 1 week ago						
An unopened UHT milk carton that expired 2 months ago						
A raw broccoli that you bought 1 week ago						
Pre-packaged hummus opened 5 days ago						
An unopened pack of dry spaghetti that expired 1 year ago						
3-day-old leftovers of a home-made chicken-curry						
An unopened can of tuna that expired 1 week ago						
A jar of green pesto that you opened 2 weeks ago						
(Reheated) leftovers of a pre-prepared veggie lasagna that you baked 4 days ago						

15. For each of the following product categories, what percentage of the edible part of what you buy or cook usually ***ends up being thrown away,*** *on average?*

	ca 0%	< 10%	< 20%	< 30%	< 40%	< 50%	> 50%	I never have this kind of food
Fresh fruit & vegetables:	O	O	O	O	O	O	O	O
Milk & dairy products:	O	O	O	O	O	O	O	O
Self-prepared meals & leftovers:	O	O	O	O	O	O	O	O

*16. In terms of behavior, how important is it for you to minimize the amount of **food that you throw away**?*

Not at all important 1	2	3	4	5	6	7	8	Extremely important 9
O	O	O	O	O	O	O	O	O

*17. Please provide your best guess of the following number: how much of the **food in households in Europe** is wasted (as a percentage of the total food bought)?*

_____%

*18. Do you have any **food intolerances, allergies or diet restrictions** for medical or non-medical reasons?*

No: O
Yes: O → if yes, which one(s)?

Intolerance or allergy to nuts	O	Diabetes	O
Gluten intolerance or allergy	O	Vegan/vegetarian	O
Lactose intolerance or allergy	O	Other:	
Intolerance or allergy to eggs	O		

*19. Have you ever experienced a serious case of **food poisoning**?*

No: O
Yes: O

If yes, did it occur in the last 12 months?

No: O
Yes: O

If yes, indicate the type of food that made you sick:
……

*20. Please tell us, in general, how willing or unwilling you are to **take risks**.*

Completely unwilling to take risks 1	2	3	4	5	6	7	8	Very willing to take risks 9
O	O	O	O	O	O	O	O	O

*21. For each of the following statements, please indicate the likelihood of **engaging in each activity**. Provide a rating from 1 to 5, using the scale: 1: Extremely unlikely to 5: Extremely likely.*

	Extremely unlikely 1	2	3	4	Extremely likely 5
Engaging in unprotected sex.	O	O	O	O	O
Not wearing a seatbelt when riding as a front-seat passenger.	O	O	O	O	O
Not wearing a helmet when riding a bicycle.	O	O	O	O	O
Exposing yourself to the sun without using sunscreen.	O	O	O	O	O
Walking home alone at night in a somewhat unsafe area.	O	O	O	O	O
Regularly eating high sugar foods.	O	O	O	O	O

Eating expired food products that still 'look okay.'	O	O	O	O	O
Buying an illegal drug for your own use.	O	O	O	O	O
Consuming five or more servings of alcohol in a single evening.	O	O	O	O	O

*22. Please indicate **your gender**:* O Male O Female

*23. Please indicate **your age**:* …….. years old

*24. Please indicate all **your nationalities***: O Swiss O Other: ……..

*25. Please indicate the highest **educational degree** that you have obtained*:

Matura/ Secondary O Bachelor O Master O PhD O Other: ……

*26. What is (are) **your field(s) of study**? (You can select multiple answers.)*

O	Architecture & Civil engineering
O	Business & Economics
O	Engineering sciences (Computer sciences, Mechanical/Electrical engineering etc.)
O	Environmental sciences
O	Other Natural sciences (Physics, Biology, Health sciences, Earth sciences…)
O	Law
O	Math
O	Medicine & Veterinary medicine
O	Political Science
O	Psychology
O	Other Arts & Humanities (Linguistics, Literature, History…)
O	Theology
O	Other: ……

*27. Had you **heard of this study** from someone who previously participated? (This will not affect your payment.)*

No: O Yes: O If yes, please explain: ……

This is the end of the questionnaire. Please, remain seated until you have finished eating the sandwich. Once you have finished it, or if you do not want to finish it, please raise your hand and wait for a staff member to come to you.

If you have any comments or questions about this study, please write them here:

Universität Zürich[UZH]

Appendix B – Exploring Attitudes Towards Food Waste

Tables

Table B-1: Summary statistics of relevant variables

Variable	N	Mean	SD	Min	Max	Description
Food Waste (15 items)	258	0.38	0.17	0.02	0.91	Is 0 if wastes very little food, 1 if wastes a lot of food, in [0;1]
FW Guess	258	0.37	0.13	0.15	0.88	Belief about the percentage of the food bought in households in Europe that is wasted, in [0; 1]
Min Waste	258	0.85	0.17	0	1	Is 1 if it is very important for the person to minimize the amount of food that she throws away, 0 otherwise, in {0, 0.125, 0.25, 0.375, … 1}
Risk General	258	0.54	0.24	0	1	Is 1 if is very willing to take risks in general, 0 otherwise, in {0, 0.125, 0.25, 0.375, … 1}
Risk Food (11 items)	258	0.64	0.17	0.13	1	Is 1 if thinks that old food is always completely safe to eat, 0 otherwise, in [0; 1]
Risk Health (9 items)	258	0.47	0.15	0.08	0.83	Is 1 if takes a lot of risks in the health/safety domain, 0 otherwise, in [0; 1]
WTP	221*	0.77	0.49	0	2	Willingness-to-pay for the second half of the sandwich, in [0; 2.2] CHF
Weight	258	12.72	24.51	0	78	Weight (in grams) of the unfinished first sandwich half, in [0; 82]

See next page for the rest of the table.

Hungry	258	5.38	1.93	1	9	Is 1 if is not at all hungry, is 9 if extremely hungry, in {1, 2, 3, … 9}
Already Eat.	258	0.63	0.48	0	1	Is 1 if has already eaten this type of sandwich in the past, 0 otherwise
Food Pois.	258	0.28	0.45	0	1	Is 1 if has already experienced a serious case of food poisoning in the past, 0 otherwise
Prep. Meals	258	4.02	1.03	1	5	Frequency of meal preparation with raw ingredients (1: Never, 2: <= 3x/month, 3: 1x or 2x/week, 4: >2x/week, 5: daily or almost daily)
Female	258	0.52	0.50	0	1	Is 1 if is a woman, 0 otherwise
Age	258	22.47	3.07	18	35	Age
Swiss	258	0.79	0.41	0	1	Is 1 if the person has the Swiss nationality, 0 if the person is a foreigner

*Notes: Summary statistics of Food Waste, Min Waste, Risk General, Risk Food and Risk Health after having normalized them so that they lie between 0 and 1. *37 individuals were not even willing to receive the second sandwich half for free.*

Table B-2: List of items to assess risk preferences in the health/safety domain

1. Engaging in unprotected sex.
2. Not wearing a seatbelt when riding as a front-seat passenger.
3. Not wearing a helmet when riding a bicycle.
4. Exposing yourself to the sun without using sunscreen.
5. Walking home alone at night in a somewhat unsafe area.
6. Regularly eating high sugar foods.
7. Eating expired food products that still 'look okay.'
8. Buying an illegal drug for your own use.
9. Consuming five or more servings of alcohol in a single evening.

Table B-3: Average of *Food Waste* across individual characteristics

	Men	**Women**	**p-value of two-sided t-tests**
$\overline{\boldsymbol{Food\ Waste}}$	0.36	0.40	0.08
Number of obs.	123	135	-
	Swiss*	**Not Swiss**	
$\overline{\boldsymbol{Food\ Waste}}$	0.36	0.45	0.001
Number of obs.	204	54	-
	Studies natural sciences**	**Does not study natural sciences**	
$\overline{\boldsymbol{Food\ Waste}}$	0.36	0.40	0.05
Number of obs.	123	135	-
	Food poisoned in the past	**Not food poisoned in the past**	
$\overline{\boldsymbol{Food\ Waste}}$	0.39	0.38	0.67
Number of obs.	72	186	-
	Prepare meals > 2x/week	**Prepare meals <= 2x/week**	
$\overline{\boldsymbol{Food\ Waste}}$	0.36	0.44	0.001
Number of obs.	186	72	-

*Notes: N=258. . Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food *This consists of all Swiss participants who also eventually have an (some) additional nationality(ies). **Natural sciences typically include: environmental sciences, physics, biology, chemistry, health sciences, earth sciences…*

Table B-4: Average value of individual traits across treatment groups

	Treatment group				
Individual characteristics	**Salient-Old**	**Salient-Fresh**	**Blind-Old**	**Blind-Fresh**	**p-value of Kruska-Wallis H test**
Food Waste	0.39	0.37	0.35	0.41	0.31
Risk General	0.5	0.56	0.57	0.52	0.27
Risk Health	0.44	0.49	0.5	0.45	0.08
Risk Food	0.64	0.65	0.67	0.61	0.13
FW Guess	0.36	0.39	0.37	0.36	0.33
Min Waste	0.86	0.87	0.84	0.84	0.91

Notes: N=258; Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise.

Table B-5: Relationship between the reported amount of food trashed and several individual traits

Dependent variable:	Food Trashed					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk General	-0.0230	-0.0128				
	(-0.64)	(-0.34)				
Risk Food			-0.176***	-0.171***		
			(-3.25)	(-3.16)		
Risk Health					-0.0288	0.00514
					(-0.49)	(0.08)
FW Guess	0.265***	0.243***	0.229***	0.207**	0.268***	0.246***
	(3.13)	(2.90)	(2.74)	(2.47)	(3.14)	(2.92)
Min Waste	-0.230***	-0.235***	-0.230***	-0.233***	-0.233***	-0.237***
	(-4.36)	(-4.51)	(-4.73)	(-4.83)	(-4.51)	(-4.65)
Hungry		0.000812		0.00133		0.000650
		(0.17)		(0.29)		(0.14)
Food Pois.		0.00714		0.00655		0.00658
		(0.38)		(0.37)		(0.35)
Female		0.0277*		0.0252		0.0295*
		(1.76)		(1.61)		(1.81)
Age		0.00138		0.000742		0.00138
		(0.53)		(0.28)		(0.54)
Constant	0.355***	0.313***	0.462***	0.434***	0.357***	0.304***
	(5.98)	(3.40)	(7.65)	(4.58)	(5.41)	(3.19)
N	258	258	258	258	258	258
R-sq.	0.153	0.163	0.196	0.204	0.153	0.163

*Notes: Coefficient estimates of linear regression models. Dependent variable: Food Trashed is in [0,1], where 1 means that the person throws a lot of food away, and 0 means that she throws very little food away. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female = 1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, and =0 if received a 1-day-old sandwich, and Salient=1 if was informed about the production date and 0 if did not know it. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-6: Relationship between reported behaviors towards old food and several individual traits

Dependent variable:	Dislikes Old Food					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk General	-0.242***	-0.243***				
	(-4.63)	(-4.59)				
Risk Food			-0.919***	-0.917***		
			(-19.72)	(-19.69)		
Risk Health					-0.314***	-0.325***
					(-4.06)	(-3.86)
FW Guess	0.282***	0.293***	0.140**	0.147**	0.317***	0.338***
	(2.97)	(3.00)	(2.20)	(2.21)	(3.18)	(3.28)
Min Waste	-0.115	-0.118	-0.131***	-0.133***	-0.153*	-0.155*
	(-1.50)	(-1.55)	(-3.04)	(-3.07)	(-1.89)	(-1.94)
Hungry		-0.00465		-0.00375		-0.00570
		(-0.68)		(-0.95)		(-0.83)
Food Pois.		0.0223		0.0125		0.0176
		(0.84)		(0.74)		(0.64)
Female		-0.00332		-0.000184		-0.0145
		(-0.14)		(-0.01)		(-0.55)
Age		0.00259		-0.00113		0.00154
		(0.68)		(-0.49)		(0.39)
Constant	0.569***	0.533***	1.066***	1.107***	0.604***	0.604***
	(6.86)	(4.29)	(19.47)	(14.94)	(6.34)	(4.46)
N	258	258	258	258	258	258
R-sq.	0.162	0.169	0.646	0.648	0.136	0.143

*Notes: Coefficient estimates of linear regression models. Dependent variable: Dislikes Old Food is in [0,1], where 1 means that the person is very unlikely to eat old food, and 0 means that she is very likely to eat old food. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female = 1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, and =0 if received a 1-day-old sandwich, and Salient=1 if was informed about the production date and 0 if did not know it. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-7: Relationship between reported food waste behaviors and several individual traits

Dependent variable:	Food Waste A					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk General	-0.173***	-0.170***				
	(-5.13)	(-4.99)				
Risk Food			-0.512***	-0.506***		
			(-13.10)	(-12.62)		
Risk Health					-0.171***	-0.153***
					(-3.17)	(-2.65)
FW Guess	0.237***	0.228***	0.172***	0.158***	0.271***	0.265***
	(3.51)	(3.37)	(3.00)	(2.74)	(3.75)	(3.66)
Min Waste	-0.132**	-0.136**	-0.146***	-0.150***	-0.158***	-0.162***
	(-2.36)	(-2.41)	(-3.44)	(-3.48)	(-2.74)	(-2.81)
Hungry		-0.000102		0.00000384		-0.00120
		(-0.02)		(0.00)		(-0.27)
Food Pois.		0.0151		0.00839		0.0108
		(0.83)		(0.59)		(0.58)
Female		0.0131		0.0184		0.0132
		(0.78)		(1.33)		(0.73)
Age		0.00420*		0.00209		0.00362
		(1.65)		(0.92)		(1.38)
Constant	0.583***	0.486***	0.839***	0.787***	0.580***	0.495***
	(9.52)	(5.75)	(16.78)	(10.88)	(8.52)	(5.33)
N	258	258	258	258	258	258
R-sq.	0.202	0.213	0.464	0.470	0.154	0.163

*Notes: Coefficient estimates of linear regression models. Dependent variable: Food Waste A is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female = 1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, and =0 if received a 1-day-old sandwich, and Salient=1 if was informed about the production date and 0 if did not know it. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-8: Relationship between reported food waste behaviors and several individual traits

Dependent variable:	Food Waste B					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk General	-0.132***	-0.128***				
	(-3.81)	(-3.68)				
Risk Food			-0.547***	-0.544***		
			(-15.93)	(-15.92)		
Risk Health					-0.171***	-0.160***
					(-3.28)	(-2.85)
FW Guess	0.273***	0.268***	0.185***	0.177***	0.293***	0.292***
	(4.02)	(3.96)	(3.31)	(3.16)	(4.11)	(4.14)
Min Waste	-0.172***	-0.177***	-0.180***	-0.183***	-0.193***	-0.196***
	(-3.16)	(-3.28)	(-5.27)	(-5.39)	(-3.48)	(-3.59)
Hungry		-0.00192		-0.00121		-0.00252
		(-0.44)		(-0.41)		(-0.56)
Food Pois.		0.0147		0.00955		0.0121
		(0.81)		(0.76)		(0.66)
Female		0.0122		0.0125		0.00751
		(0.79)		(1.09)		(0.45)
Age		0.00199		-0.000196		0.00146
		(0.85)		(-0.12)		(0.62)
Constant	0.462***	0.423***	0.764***	0.770***	0.481***	0.454***
	(7.42)	(4.80)	(18.28)	(12.96)	(6.83)	(4.72)
N	258	258	258	258	258	258
R-sq.	0.205	0.212	0.586	0.589	0.188	0.193

*Notes: Coefficient estimates of linear regression models. Dependent variable: Food Waste B is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, and Risk Food and Risk Health are in [0,1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female = 1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, and =0 if received a 1-day-old sandwich, and Salient=1 if was informed about the production date and 0 if did not know it. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-9: Coefficients of partial determination based on the regressions of *Food Trashed* and *Dislikes Old Food*, respectively, on individual traits

Dependent variable:	Food Trashed			Dislikes Old Food		
	(1)	(2)	(3)	(4)	(5)	(6)
Risk Food	0.054***	-	-	0.62***	-	-
Risk General	-	0.002	-	-	0.092***	-
Risk Health	-	-	0.001	-	-	0.066***
FW Guess	0.05***	0.061***	0.064***	0.02**	0.032***	0.04***
Min Waste	0.097***	0.092***	0.095***	0.034***	0.011*	0.02**

Notes: N=258. Coefficients of partial determination of regressing Food Trashed and Dislikes Old Food on FW Guess and Min Waste, as well as on Risk Food (col. 1 and 4), Risk General (col. 2 and 5) or Risk Health (col. 3 and 6). Dependent variable: col. 1 to 3: Food Trashed is in [0,1], where 1 means that the person throw a lot of food away, and 0 means that she throws very little food away; col. 4 to 6: Dislikes Old Food is in [0,1], where 1 means that the person is very unlikely to eat old food, and 0 means that she is very likely to eat old food. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, Risk Food and Risk Health are in [0.1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food- or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise. $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-10: Coefficients of partial determination based on the regressions of alternative measures of food waste behaviors on individual traits

Dependent variable:	Food Waste A			Food Waste B		
	(1)	(2)	(3)	(4)	(5)	(6)
Risk Food	0.401***	-	-	0.518***	-	-
Risk General	-	0.10***	-	-	0.065***	-
Risk Health	-	-	0.042***	-	-	0.047***
FW Guess	0.036***	0.044***	0.056***	0.062***	0.067***	0.077***
Min Waste	0.055***	0.03***	0.041***	0.114***	0.057***	0.07***

Notes: N=258. Coefficients of partial determination of regressing Food Waste A (col. 1 to 3) and Food Waste B (col. 4 to 6) on FW Guess and Min Waste, as well as on Risk Food (col. 1 and 4), Risk General (col. 2 and 5) or Risk Health (col. 3 and 6). Dependent variable: Food Waste A and Food Waste B are in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food. Independent variables: Risk General is in {0, 0.125, 0.25, 0.375, ... 1}, Risk Food and Risk Health are in [0.1], where 0 indicates aversion to risk taking and 1 indicates love for risk taking in general, in the food- or in the health/safety domain, respectively; FW Guess is in [0,1] and represents a person's belief about how much of the food in households in Europe is wasted (as a percentage of the total food bought); Min Waste is in {0, 0.125, 0.25, 0.375, ... 1} where 1 means that it is very important for the person to minimize food waste, 0 otherwise. $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-11: Relationship between individuals' willingness-to-pay and reported food waste behaviors (Tobit model)

Dependent variable:	**Willingness-to-pay for the 2nd half**	
	(1)	**(2)**
Food Waste	-0.0728	-0.00403
	(-0.34)	(-0.02)
Already Eat.		0.0826
		(1.03)
Hungry		0.0535**
		(2.42)
Food Pois.		0.0254
		(0.29)
Female		-0.0804
		(-0.99)
Age		0.0000627
		(0.00)
Constant	0.874***	0.535
	(7.65)	(1.58)
N	221	221
Pseu. R-sq.	0.008	0.028

*Notes: Estimated marginal effects of Tobit models. Dependent variables: WTP is in [0; 2.2] CHF (the 37 observations with a negative WTP are excluded from the analysis). Independent variables: Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food; Already Eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female =1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, =0 if received a 1-day-old sandwich and Salient=1 if was informed about the production date and 0 if did not know it. McFadden Pseudo R2 are provided. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-12: Relationship between individuals' willingness-to-pay for the second sandwich half and alternative measures of reported food waste behaviors (Hurdle model)

Dependent variable:	WTP weakly pos. (1/0)	WTP	WTP weakly pos. (1/0)	WTP	WTP weakly pos. (1/0)	WTP	WTP weakly pos. (1/0)	WTP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Food Waste A	-0.219	-0.175	0.0848	-0.125				
	(-0.32)	(-1.02)	(0.12)	(-0.72)				
Food Waste B					-0.261	-0.0908	0.0728	-0.0292
					(-0.39)	(-0.51)	(0.10)	(-0.16)
Already Eat.			0.417**	0.110*			0.417**	0.113*
			(2.01)	(1.77)			(2.01)	(1.81)
Hungry			0.151***	0.0211			0.151***	0.0209
			(3.04)	(1.40)			(3.04)	(1.39)
Food Pois.			-0.0106	0.0682			-0.0109	0.0662
			(-0.05)	(1.14)			(-0.05)	(1.10)
Female			-0.426**	-0.0492			-0.425**	-0.0560
			(-2.07)	(-0.81)			(-2.06)	(-0.93)
Age			-0.0449	-0.00213			-0.0446	-0.00249
			(-1.49)	(-0.23)			(-1.49)	(-0.27)
Constant	1.059***	1.006***	1.135	0.854***	1.047***	0.955***	1.141	0.817***
	(2.82)	(10.04)	(1.47)	(3.64)	(3.41)	(10.95)	(1.51)	(3.56)
N	258	183	258	183	258	183	258	183
Pseu. R-sq.	0.013		0.101		0.014		0.101	

*Notes: Columns 1, 3, 5 and 7: Estimated marginal effects of Probit models. Columns 2, 4, 6 and 8: Estimated marginal effects of truncated regressions at 0. Columns 1 to 8 represent four hurdle models. Dependent variable: Col. 1, 3, 5 and 7: WTP weakly pos.=1 if WTP>=0, and WTP weakly pos.=0 if WTP<0; Col. 2, 4, 6 and 8: WTP in [0; 2.2] CHF. Independent variables: Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food; Already Eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female =1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, =0 if received a 1-day-old sandwich and Salient=1 if was informed about the production date and 0 if did not know it. McFadden Pseudo R2 are provided in col. 1, 3, 5, and 7. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-13: Relationship between the decision to eat entirely the first sandwich half and measures of reported food waste behaviors

Dependent variable:	First half entirely eaten (0/1)					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Food Waste	-0.517***	-0.384***				
	(-3.35)	(-2.64)				
Food Waste A			-0.517***	-0.346**		
			(-2.89)	(-2.21)		
Food Waste B					-0.519***	-0.336**
					(-2.82)	(-2.11)
Already Eat.		0.137***		0.138***		0.138***
		(2.69)		(3.09)		(3.09)
Hungry		0.0624***		0.0551***		0.0549***
		(4.45)		(4.98)		(4.96)
Food Pois.		-0.00618		-0.0120		-0.00891
		(-0.12)		(-0.24)		(-0.18)
Female		-0.218***		-0.209***		-0.210***
		(-4.55)		(-4.62)		(-4.64)
Age		-0.00624		-0.00644		-0.00760
		(-0.71)		(-0.82)		(-0.99)
N	258	258	258	258	258	258
R-sq.	0.053	0.227	0.038	0.220	0.037	0.218

*Notes: Columns 1 and 2: Coefficient estimates of linear regression models, Columns 3 to 6: Estimated marginal effects of Probit models. Dependent variables: First half entirely eaten =1 if the subject finished to eat his first sandwich half, 0 otherwise. Independent variables: Food Waste is in [0,1], where 1 means that the person wastes a lot of food, and 0 means that she wastes very little food; Already Eat.=1 if has already eaten this type of sandwich in the past, 0 otherwise; Hungry is in {1;2;…;9} where 1 means "Not at all hungry" and 9 means "Extremely hungry"; Food Pois.=1 if has ever experienced a serious case of food poisoning, 0 otherwise; Female =1 if is a woman, 0 otherwise. All models control for the treatment conditions (Old, Salient and Salient*Old), where Old=1 if received a 7-day-old sandwich, =0 if received a 1-day-old sandwich and Salient=1 if was informed about the production date and 0 if did not know it. McFadden Pseudo R2 are provided in col. 3 to 6. Robust standard errors, t-statistics in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table B-14: Comparison of the perceived and actual safety of old food items

		How **safe** do you think it is to eat this item?			
nr	**Food item description**	**Not at all safe**	**Moderately safe**	**Completely safe**	**Actual safety***
1	A delivered (reheated) pizza with vegetables and meat that has been sitting in your fridge for 3 days	0.1	0.64	0.26	*Microbiological analyses required*
2	3-day-old leftovers of a home-made chicken-curry	0.09	0.54	0.37	*Microbiological analyses required*
3	(Reheated) leftovers of a pre-prepared veggie lasagna that you baked 4 days ago	0.11	0.58	0.31	*Microbiological analyses required*
4	An unopened plain (i.e., unflavored) yoghurt that expired 10 days ago	0.19	0.47	0.34	**Completely safe** (*SL1-2, ND, SCA, FS)*
5	An unopened UHT milk carton that expired 2 months ago	0.38	0.41	0.22	**Completely safe** (*SL1-2, ND, SCA, FS)*
6	An unopened pack of dry spaghetti that expired 1 year ago	0.05	0.26	0.69	**Completely safe** (*SL1, ND, FS)*
7	An unopened can of tuna that expired 1 week ago	0.14	0.27	0.59	**Completely safe** (*SL1-2, ND, FS)*
8	A raw broccoli that you bought 1 week ago	0.02	0.27	0.71	**Completely safe** (*SL1*[95]*, FS)*
9	Hard boiled eggs that you cooked 1 week ago	0.24	0.36	0.4	**Moderately safe** (*SL1, ND, SCA, FS)*

See next page for the rest of the table.

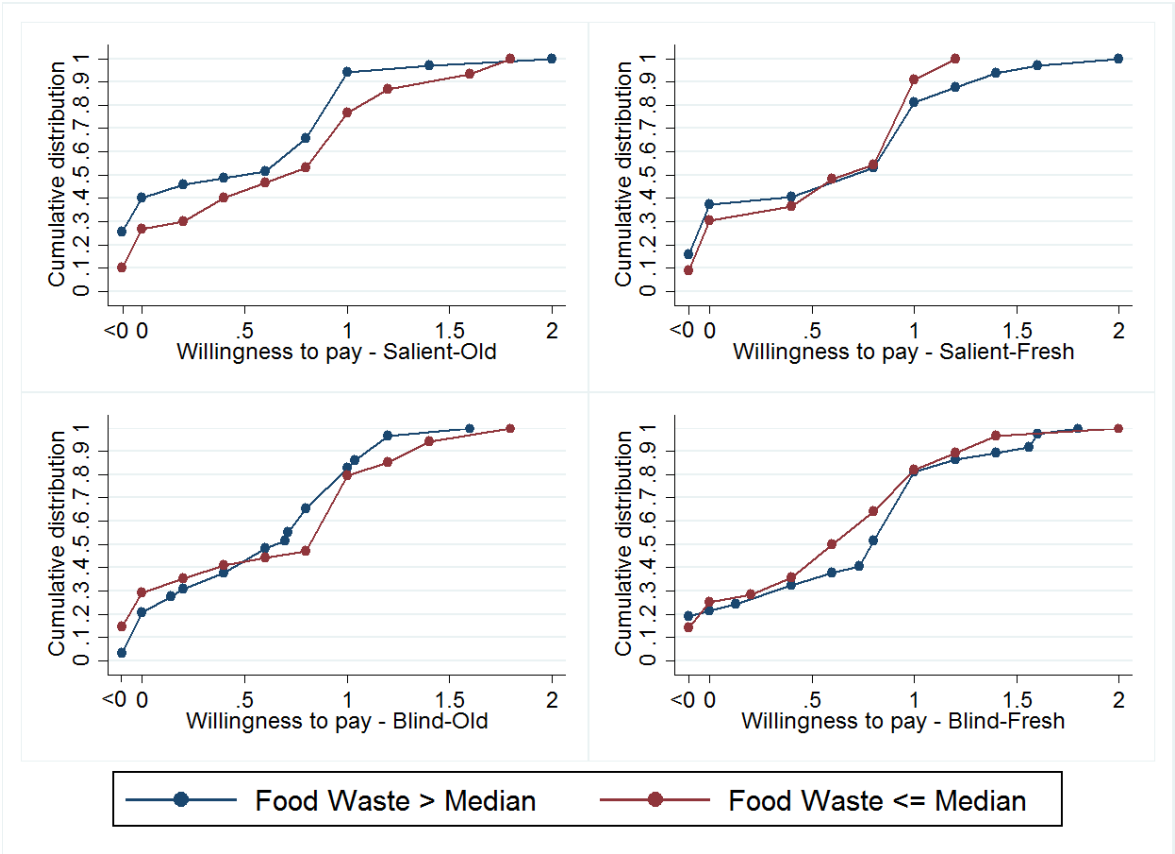
[95] The Swiss law does not require any date label on fresh vegetables and fruits, as their edibility can be easily evaluated by a sensory evaluation. As we specify in the question's statement that every item is assumed to look and smell fine, this raw broccoli is perfectly safe to eat.

10	Pre-packaged hummus opened 5 days ago	0.07	0.58	0.35	*Microbiological analyses required*
11	A jar of green pesto that you opened 2 weeks ago	0.16	0.51	0.33	*Microbiological analyses required*

Notes: N_{nr}=258, where nr *represents the item number. However, participants missed reporting their likelihood to eat some food items (implying: N_3= N_{10}=257, N_6=256), and their perception of their safety (implying: N_3=N_9= N_{10}=257, N_6=256). Subjects were asked to assume that these items had been stored in ideal conditions since the purchase date, and that their smell and appearance was fine when opening or inspecting them.*
**The actual safety of these items is evaluated based on what the Swiss law (SL) prescribes (SL1 refers to the* Hygiene ordinance *(2016), SL2 refers to the* Ordinance on the labeling and advertising of food stuff *(2005)), the report from a Swiss consumer association (FRC, 2012) and the emails exchanged with its author (SCA), the opinion of a food scientist working for one of the biggest Swiss supermarket (FS), as well as an article published by a nutritionist dietician (ND) in a recognized French journal specialized in nutrition and health (Rivet-Bonjean, 2018).*

Figures

Figure B-1: Cumulative distributions of subjects' willingness-to-pay for the second sandwich half with respect to their reported food waste behaviors, by treatment group



Notes: N=258 with $N_{Salient-Old}$*=65,* $N_{Salient-Fresh}$*=65,* $N_{Blind-Old}$*=65 and* $N_{Blind-Fresh}$*=63. Food Waste is in [0,1],* where *1 means that the person wastes a lot of food, and 0 means that she wastes very little food.*

Appendix C – Psychological Opportunity Costs: The Effect of Opportunity Costs on Post-Choice Utility

Tables

Table C-1: Description of the goods in each good category

Good category	Description of the six goods in each category
Individual portion (175ml) of a Mövenpick ice cream	Stracciatella, Swiss chocolate, Vanilla dream, Caramelita, Pistachio, Maple walnut.
Voucher for an activity	Bungee jumping in the Swiss mountains (30 min), Round flight in the Swiss mountains (30 min), Chocolate workshop (around 4 hours), Sushi lesson (around 2.5 hours), Thai massage (1 hour), Rafting tour in Switzerland (around 4 hours).
Caran d'Ache ball pen (blue ink)	Black, dark blue, dark green, light blue, red, white.
Short educative animations (ca 3min)*	*7 tips to wake up without coffee, How is your phone changing you, How to increase your productivity, How much sleep do you actually need, 5 ways social media is changing your brain, The 9 best scientific study tips.*
Voucher for a dish in a local restaurant	Beef with thai basil leaves, Veggie dumplings with salad, Gnocchi with tomato sauce, basil, grana Padano cheese and fior di latte, Lamb with vegetables tajine, Spaghetti with tomatoes, basil, garlic and olive oil, Braised tofu with vegetables in clay pot.

** These animations are free of access on www.youtube.com. We downloaded them and cut them such as to remove the advertisement usually located towards the end.*

Table C-2: Pictures of the 30 goods presented to the subjects

Ice Cream	Pen	Dish	Movie	Activity
				
				
				
				
				
				

Table C-3: Measure of decisional difficulty in the on-line survey

The ten sentences below describe behaviors and thoughts that one typically experiences when having difficulty making up ones' mind. Subjects had to evaluate to what extent each sentence apply to their case on a scale from 1: *Doesn't apply at all* to *7: Does perfectly apply.*

Decision Difficulty scale

1	When I am in the car listening to the radio, I often check other stations to see if something better is playing, even if I'm relatively satisfied with what I'm listening to.
2	I often change my mind several times before making a decision.
3	Even if I see a choice I really like, I have a hard time making the decision if I do not have a chance to check out other possible options.
4	Whenever I'm faced with a choice, I try to imagine what all the other possibilities are, even ones that aren't present at the moment.
5	No matter how satisfied I am with my job, it's only right for me to be on the lookout for better opportunities.
6	I always keep my options open so I will not miss the next best choice available in life.
7	I often fantasize about living in ways that are quite different from my actual life.
8	Whenever I made a choice, I'm curious about what would have happened if I had chosen differently.
9	I often wonder why decisions can't be easier.
10	I am usually worried about making a wrong decision.

Table C-4: Summary statistics of relevant variables

Variable	N	Mean	SD	Min	Max	Description
Stage 1 Happ.	630	78	21.9	1	100	Hypothetical happiness level reported in the on-line survey with the good that was chosen/allocated in the laboratory. 1 means *Not at all happy* and 100 means *Very happy*, in {1,2,3,…,100}
Post-Choice Happ.	630	70.7	22.6	1	100	Effective happiness level with the chosen/allocated good in the laboratory after the decision/allocation occurred. 1 means *Not at all happy* and 100 means *Very happy*, in {1,2,3,…,100}
Post-Consumption Happ.	182	66.8	24.6	1	100	Effective happiness level with the chosen/allocated good in the laboratory after consumption occurred. 1 means *Not at all happy* and 100 means *Very happy*, in {1,2,3,…,100}

Table C-5: Relationship between post-choice happiness and the treatment condition *Choice* versus *No-Choice*

	(1)	(2)	(3)	(4)
Dependent variable:		**Post-Choice Happ.**		
Model:	**OLS**	**OLS**	**Tobit**	**Tobit**
***dummy*Choice**	-4.111***	-4.164***	-4.574**	-4.960***
	(1.566)	(1.589)	(1.775)	(1.822)
Stage 1 Happ.	0.663***	0.576***	0.699***	0.584***
	(0.0341)	(0.0392)	(0.0387)	(0.0434)
***dummy*Pen**		-5.564**		-7.095***
		(2.219)		(2.241)
***dummy*Dish**		0.581		-0.625
		(1.561)		(1.744)
***dummy*Movie**		-6.237***		-7.193***
		(1.966)		(2.212)
***dummy*Activity**		7.961***		9.633***
		(1.545)		(1.927)
Constant	21.99***	29.45***	21.03***	31.20***
	(3.046)	(3.694)	(3.416)	(4.132)
Observations	630	630	608	608
R^2	0.419	0.463		

*Notes: Coefficient estimates of linear regression models (columns 1 and 2) and of Tobit models (columns 3 and 4). The observations are aggregated across all good categories together. Observations qualified as "non-eligible" (i.e., cases where individuals had already watched at least one of the two movie clips, could not eat ice creams or who misreported their happiness level) are added back to the data set in columns 1 and 2. Dependent variable: Post-Choice Happ. is the happiness level with the chosen/allocated option in the laboratory (pre-consumption) and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyChoice=1 if is in the Choice condition, and 0 if is in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyPen, dummyDish, dummyMovie and dummyActivity are the good categories fixed effects and are in {0,1} (e.g. dummyPen=1 if the post-choice happiness is reported for the Pen good category, and 0 otherwise). The Ice Cream good category represents the reference category. Standard errors clustered at the individual level in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table C-6: Relationship between post-choice happiness and the choice conditions *Low POCs* versus *High POCs*

	(1)	(2)	(3)	(4)
Dependent variable:		**Post-Choice Happ.**		
Model:	**OLS**	**OLS**	**Tobit**	**Tobit**
***dummy*NoChoice**	5.185***	5.076***	5.593***	5.730***
	(1.770)	(1.809)	(2.004)	(2.063)
***dummy*HighPOCs**	2.146	1.823	2.021	1.535
	(1.608)	(1.550)	(1.704)	(1.633)
Stage 1 Happ.	0.664***	0.578***	0.700***	0.586***
	(0.0337)	(0.0389)	(0.0385)	(0.0433)
***dummy*Pen**		-5.554**		-7.047***
		(2.207)		(2.235)
***dummy*Dish**		0.568		-0.604
		(1.575)		(1.757)
***dummy*Movie**		-6.171***		-7.091***
		(1.967)		(2.215)
***dummy*Activity**		7.942***		9.638***
		(1.554)		(1.932)
Constant	16.69***	24.25***	15.29***	25.29***
	(3.027)	(3.635)	(3.469)	(4.104)
Observations	630	630	608	608
R^2	0.420	0.464		

*Notes: Coefficient estimates of linear regression models (columns 1 and 2) and of Tobit models (columns 3 and 4). The observations are aggregated across all good categories together. Observations qualified as "non-eligible" (i.e., cases where individuals had already watched at least one of the two movie clips, could not eat ice creams or who misreported their happiness level) are added back to the data set in columns 1 and 2. Dependent variable: Post-Choice Happ. is the happiness level with the chosen/allocated option in the laboratory and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyNoChoice=1 if is in the No-Choice condition, and 0 if is in the Choice condition; dummyHighPOCs=1 if is in the High POCs condition, and 0 if is in the Low POCs or in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyPen, dummyDish, dummyMovie and dummyActivity are the good categories fixed effects and are in {0,1} (e.g. dummyPen=1 if the post-choice happiness is reported for the Pen good category, and 0 otherwise). The Ice Cream good category represents the reference category. Standard errors clustered at the individual level in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table C-7: Relationship between post-consumption happiness and the treatment conditions *Choice* versus *No-Choice*

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Post-Consumption Happ.**					
Model:	**OLS**	**OLS**	**OLS**	**OLS**	**Tobit**	**Tobit**
***dummy*Choice**	0.116	-0.647	-0.213	-1.713	2.157	0.402
	(3.436)	(3.462)	(3.387)	(3.403)	(3.751)	(3.742)
Stage 1 Happ.	0.542***	0.427***	0.533***	0.420***	0.607***	0.483***
	(0.0835)	(0.0855)	(0.0866)	(0.0876)	(0.0944)	(0.0928)
***dummy*Movie**		-19.07***		-18.99***		-20.51***
		(2.463)		(2.400)		(2.695)
Constant	23.43***	43.89***	24.15***	44.36***	18.95**	41.08***
	(7.597)	(8.118)	(7.836)	(8.308)	(8.263)	(8.444)
Observations	197	197	202	202	182	182
R^2	0.197	0.332	0.190	0.331		

*Notes: Coefficient estimates of linear regression models (columns 1 to 4) and of Tobit models (columns 5 and 6). The observations are aggregated across all good categories together. Observations qualified as "non-eligible" (i.e., cases where individuals had already watched at least one of the two movie clips, or could not eat ice creams) are added back to the data set in columns 1 and 2. Individuals who were qualified as "non-eligible" or/and who expressed regret with their choice are added back to the data set in columns 3 and 4. Dependent variable: Post-Consumption Happ. is the happiness level with the chosen/allocated option after consumption occurred and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyChoice=1 if is in the Choice condition, and 0 if is in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyMovie=1 if (post-consumption) happiness is reported for the Movie good category, and 0 if is reported for the Ice Cream good category. Standard errors clustered at the individual level in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table C-8: Relationship between post-consumption happiness and the choice conditions *Low POCs* versus *High POCs*

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Post-Consumption Happ.**					
Model:	**OLS**	**OLS**	**OLS**	**OLS**	**Tobit**	**Tobit**
***dummy*NoChoice**	-1.242	-0.421	0.0312	1.317	-2.185	-0.764
	(3.761)	(3.690)	(3.825)	(3.713)	(4.076)	(3.952)
***dummy*HighPOCs**	-2.701	-2.561	-0.406	-0.888	-0.0675	-0.895
	(3.742)	(3.182)	(3.521)	(2.930)	(4.054)	(3.389)
Stage 1 Happ.	0.538***	0.423***	0.532***	0.419***	0.607***	0.481***
	(0.0841)	(0.0861)	(0.0864)	(0.0875)	(0.0944)	(0.0932)
***dummy*Movie**		-19.06***		-19.00***		-20.54***
		(2.470)		(2.399)		(2.686)
Constant	25.00***	44.61***	24.16***	43.15***	21.14***	41.99***
	(6.619)	(6.887)	(6.842)	(7.074)	(7.145)	(7.132)
Observations	197	197	202	202	182	182
R^2	0.199	0.334	0.190	0.331		

*Notes: Coefficient estimates of linear regression models (columns 1 to 4) and of Tobit models (columns 5 and 6). The observations are aggregated across all good categories together. Observations qualified as "non-eligible" (i.e., cases where individuals had already watched at least one of the two movie clips, or could not eat ice creams) are added back to the data set in columns 1 and 2. Individuals who were qualified as "non-eligible" or/and who expressed regret with their choice are added back to the data set in columns 3 and 4. Dependent variable: Post-Consumption Happ. is the happiness level with the chosen/allocated option after consumption occurred and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyNoChoice=1 if is in the No-Choice condition, and 0 if is in the Choice condition; dummyHighPOCs=1 if is in the High POCs condition, and 0 if is in the Low POCs or in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; dummyMovie=1 if (post-consumption) happiness is reported for the Movie good category, and 0 if is reported for the Ice Cream good category. Standard errors clustered at the individual level in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table C-9: Relationship between post-choice happiness and the treatment condition *Choice* versus *No-Choice*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:				**Post-Choice Happ.**			
Good category:	**All**	**All**	**Ice Cream**	**Pen**	**Dish**	**Movie**	**Activity**
***dummy*Choice**	-3.109*	-3.413*	-4.795	2.215	-5.484*	-7.859**	-1.267
	(1.676)	(1.777)	(2.993)	(3.657)	(2.838)	(3.592)	(2.709)
Stage 1 Happ.	0.652***	0.552***	0.603***	0.453***	0.529***	0.651***	0.665***
	(0.0360)	(0.0410)	(0.0885)	(0.0738)	(0.0852)	(0.0900)	(0.121)
Choice Diff.	0.132	0.136	0.0597	0.109	0.167	0.0311	0.282
	(0.0878)	(0.0923)	(0.150)	(0.266)	(0.155)	(0.149)	(0.180)
Choice Diff. x	-0.0484	-0.0627	-0.0113	-0.0894	0.0186	0.0712	-0.293
***dummy*Choice**	(0.122)	(0.124)	(0.196)	(0.316)	(0.195)	(0.203)	(0.203)
***dummy*Pen**		-6.765***					
		(2.157)					
***dummy*Dish**		-0.257					
		(1.564)					
***dummy*Movie**		-6.534***					
		(2.085)					
***dummy*Activity**		7.437***					
		(1.501)					
Constant	22.45***	31.56***	28.38***	26.99***	34.71***	20.87**	26.77**
	(3.261)	(4.010)	(8.271)	(5.569)	(7.781)	(8.198)	(12.04)
Observations	608	608	121	126	126	109	126
R^2	0.407	0.454	0.381	0.250	0.292	0.425	0.274
SE	Clustered	Clustered	Robust	Robust	Robust	Robust	Robust

*Notes: Coefficient estimates of linear regression models. Dependent variable: Post-Choice Happ. is the happiness level with the chosen/allocated option in the laboratory (pre-consumption) and is in {1,2,3…,100} where 1 means "Not at all happy", and 100 means "Very happy". Independent variables: dummyChoice=1 if is in the Choice condition, and 0 if is in the No-Choice condition; Stage 1 Happ. is the (choiceless) happiness level (reported in the on-line survey) with the chosen/allocated option in the laboratory; Choice Diff. is mean-centered and is in [-38.45, 35.55] where higher values represent a higher tendency to experience choice difficulty; dummyPen, dummyDish, dummyMovie and dummyActivity are the good categories fixed effects and are in {0,1} (e.g. dummyPen=1 if the post-choice happiness is reported for the Pen good category, and 0 otherwise). The Ice Cream good category represents the reference category. Standard errors are clustered at the individual level in columns 1 and 2. Standard errors in parentheses, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Appendix D – Immoral Labor Markets

Proofs

Lemma. *For all* $j \in J^{IM}$, $w^*(j)$ *exists, is unique and is in* $(\underline{u}+c, \infty)$. *For all* $j \in J \setminus J^{IM}$, $w^*(j) = \underline{u}+c$.

Proof. Suppose $j \in J^{IM}$. *Existence*: Define $f(w,j) = S(w,j) - D(w,j)$. Note that $f(\underline{u}+c, j) = 0 - (+) < 0$, $lim_{w\to\infty} f(w,j) = 1 - 0 = 1$ and $f(w,j)$ is continuous in w. By the intermediate value theorem there exists $w^*(j) \in (\underline{u}+c, \infty)$ such that $f(w^*(j), j) = 0$. *Uniqueness*: Follows from $f(w,j)$ being strictly increasing in w on $[0,\infty)$. Suppose $j \in J \setminus J^{IM}$. Then, $S(w,j) = \begin{cases} 0, & w < \underline{u}+c \\ [0,1], & w = \underline{u}+c \\ 1, & w > \underline{u}+c \end{cases}$. Note that for any $< \underline{u}+c$, we have $D(w,j) > 0$ but $S(w,j) = 0$, and for any $w > \underline{u}+c$ we have $D(w,j) < 1$, but $S(w,j) = 1$. For $w = \underline{u}+c$, $S(w,j) = [0,1]$ and $D(w,j) \in [0,1]$, so $D(w,j) \in S(w,j)$.

Proposition 1. *For all* $j, j' \in J$ *with* $I(j) < I(j')$, $w^*(j) < w^*(j')$.

Proof. $w^*(j) < w^*(j')$: Suppose $I(j) = 0$. Then, $w^*(j) = \underline{u}+c$ and $w^*(j') > \underline{u}+c$ (see Lemma). Suppose $I(j) > 0$. Suppose that $w^*(j) \geq w^*(j')$. Then $S(w^*(j), j) > S(w^*(j'), j')$ and $D(w^*(j), j) \leq D(w^*(j'), j)$ because F and $-D$ are strictly increasing in w on $[0,\infty)$, $I(j) < I(j')$ and therefore $D(w, j') \geq D(w, j)$ for all w. So $S(w^*(j), j) - S(w^*(j'), j') + D(w^*(j'), j') - D(w^*(j), j) > 0$, a contradiction to the definition of $w^*(j)$ and $w^*(j')$.

Proposition 2. *For all* $j \in J^{IM}$, *worker* i *is hired iff* $\theta_i \leq \frac{w^*(j)-\underline{u}-c}{I(j)} \equiv \underline{\theta}(j) \in \mathbb{R}_{>0}$.

Proof. $\underline{\theta}(j) \in \mathbb{R}_{>0}$: Follows directly from $w^*(j) > \underline{u}+c$ (see Lemma).

Proposition 3. For all $j, j' \in J$ with $I(j) < I(j')$, there exists $\tilde{\theta}(j,j') \in \mathbb{R}_{>0}$ such that $U_i(j', w^*(j')) > U_i(j, w^*(j))$ iff $\theta_i < \tilde{\theta}(j,j')$.

Proof. Suppose $I(j) = 0$. Then $U_i(j, w^*(j)) = U_i(j, \underline{u} + c) = \underline{u}$ (see Lemma). Note that $U_i(accept|j', w^*(j')) > \underline{u}$ iff $\theta_i < \frac{w^*(j') - \underline{u} - c}{I(j')} = \underline{\theta}(j')$, so $U_i(j', w^*(j')) = U_i(accept|j', w^*(j')) > \underline{u} = U_i(j, w^*(j))$ iff $\theta_i < \underline{\theta}(j') \equiv \tilde{\theta}(j,j')$. Note that $\tilde{\theta}(j,j') > 0$ by Proposition 2.

Suppose $I(j) > 0$. Note that $U_i(j', w^*(j')) > U_i(j, w^*(j)) = max\{U_i(accept|j, w^*(j)), \underline{u}\}$ if and only if:

i) $U_i(accept|j', w^*(j')) > U_i(accept|j, w^*(j))$

ii) $U_i(accept|j', w^*(j')) > \underline{u}$

Inequality i) holds iff $\theta_i < \frac{w^*(j') - w^*(j)}{I(j') - I(j)}$ and inequality ii) holds iif $\theta_i < \underline{\theta}(j')$. Therefore, we have $\tilde{\theta}(j,j') = min\left\{\frac{w^*(j') - w^*(j)}{I(j') - I(j)}, \underline{\theta}(j')\right\}$. Note that $\tilde{\theta}(j,j') > 0$ (see Proposition 1 and Proposition 2).

Proposition 4. *For all $j \in J^{IM}$ and $F, G \in \mathcal{F}_\theta$ with F strong first-order stochastically dominating G, there exists $\hat{\theta}(j,F) \in \mathbb{R}_{>0}$ such that $U_i(j, w^*(j,F)) > U_i(j, w^*(j,G))$ iff $\theta_i < \hat{\theta}(j,F)$.*

Proof. First, we will proof that $w^*(j,G) \leq w^*(j,F)$. Suppose not, then $w^*(j,G) > w^*(j,F)$. But then $S(w^*(j,G), j, G) > S(w^*(j,F), j, F)$ and $D(w^*(j,G), j) < D(w^*(j,F), j)$ because F, G and $-D$ are strictly increasing in w on $[0, \infty)$, and F first-order stochastically dominates G. But then $S(w^*(j,G), j, G) - S(w^*(j,F), j, F) + D(w^*(j,F), j) - D(w^*(j,G), j) > 0$, a contradiction to the definition of w^*.

Second, define $\hat{\theta}(j,F) = \frac{w^*(j,F) - \underline{u} - c}{I(j)} \equiv \underline{\theta}(j,F)$, and note that under this definition $w^*(j,F) - c - S(j) * \theta_i > \underline{u}$ iff $\theta_i < \hat{\theta}(j,F)$. ($\hat{\theta}(j,F) > 0$ follows from Proposition 2.)

To finish the proof, note that $U_i\big(j, w^*(j,F)\big) = w^*(j,F) - c - I(j) * \theta_i > \max\{\underline{u}, w^*(j,G) - c - I(j) * \theta_i\} = U_i\big(j, w^*(j,G)\big)$ for all $\theta_i < \hat{\theta}(j,F)$, and $U_i\big(j, w^*(j,F)\big) = \underline{u} = U_i\big(j, w^*(j,G)\big)$ for all $\theta_i \geq \hat{\theta}(j,F)$.

Alternative interpretation model

The results in section 5.3 also apply for a context with 2 jobs, a neutral job j^N ($I(j^N) = 0$) and an immoral job $j^{IM} \in J^{IM}$ ($I(j^{IM}) > 0$). In the following, we show that, under some assumptions, labor demand and labor supply in this setting correspond to their counterparts in section 5.3. Therefore, all results derived in section 5.3 also hold in this context.

Labor supply: Labor supply consists of an interval of workers, $i \in [0,1]$. As in section 5.3, we assume that the utility of accepting job j of a worker of type i is given by:

$$U_i(j, w(j)) = w(j) - c - I(j) * \theta_i,$$

where the parameter θ_i is distributed according to a distribution with cdf $F \in \mathcal{F}_\theta$. For all $F \in \mathcal{F}_\theta$, F is continuous, strictly increasing on $[0,1)$, and with $F(0) = 0$. Every worker with $\theta_i \leq \frac{w(j^{IM}) - w(j^N)}{I(j^{IM})}$ accepts the immoral job. Define $w = w(j^{IM}) - w(j^N)$, the labor supply for the immoral job is then given by $S(w, j^{IM}) = F(\frac{w}{I(j^{IM})})$. Note that the labor supply corresponds to the labor supply in section 5.3 with $\underline{u} = c = 0$. The only difference lies in the interpretation: w now measures the immorality premium ($w(j^{IM}) - w(j^N)$), "accept" refers to accepting the immoral job and "refuse" refers to accepting the neutral job.

Labor demand: Labor demand is represented by an interval of firms, $k \in [0,1]$. Each firm can produce a neutral or an immoral product. Firms that produce immoral products offer immoral jobs, firms that produce neutral products offer neutral jobs. Firms profits are:

$$\pi_k(j, w) = a_k(j) - w(j)$$

where $a_k(j)$ measures firm k's earnings by producing good j. Firms offer an immoral job iff $\Delta a_k(j^{IM}) = a_k(j^{IM}) - a_k(j^N) \geq w$. We assume that $\Delta a_k(j^{IM})$ is distributed according to a distribution with cdf $G_{j^{IM}}$. An increase in immorality of the job does not decrease firms

earnings,[96] that is, i) $G_j(0) = 0$ for all $j \in J^{IM}$, and ii) for all $j, j' \in J^{IM}$ with $I(j') > I(j)$ and all $x \in \mathbb{R}$, $G_{j'}(x) \leq G_j(x)$. In addition, $G_{j^{IM}}$ is continuous and strictly increasing on $[0, \infty)$. The labor demand function for the immoral job is then given by $D(w, j^{IM}) = 1 - G_{j^{IM}}(w)$. Note that D is continuous in w, strictly decreasing in w on $[0, \infty)$, with $lim_{w \to \infty} D(w, j^{IM}) = 0$ and $D(w, j^{IM}) = 1$ for $w \leq 0$. In addition, $I(j') > I(j)$ implies $D(w, j') \geq D(w, j)$ for all $w \in \mathbb{R}$. Note that the labor demand corresponds to the labor demand in section 5.3. The only difference lies in the interpretation: $D(w, j)$ now measures the share of firms that offer the immoral job j, while $1 - D(w, j)$ measures the share of firms that offer the neutral job j^N.

The equilibrium wage, $w^*(j^{IM})$, is implicitly defined by $S(w^*(j^{IM}), j^{IM}) - D(w^*(j^{IM}), j^{IM}) = 0$.[97] As both labor demand and labor supply correspond to their counterparts in section 5.3, the Lemma and Proposition 1 to 4 (with $j \in J^{IM}$) apply. In particular, $w^*(j^{IM})$ is strictly positive (Lemma), so there is an immorality premium, and this immorality premium is increasing in the immorality of j^{IM}, $I(j^{IM})$ (Proposition 1). The immoral types sort into accepting the immoral jobs while the moral types sort into accepting the neutral jobs (Proposition 2).

[96] One interpretation is, for example, that $I(j)$ measures negative externalities in production. Avoiding these externalities is costly, decreasing the immorality therefore increases production costs (see also Rosen, 1986).

[97] Note that market clearance for the immoral job market implies market clearance for the neutral job market, $\left(1 - S(w^*(j^{IM}), j^{IM})\right) - \left(1 - D(w^*(j^{IM}), j^{IM})\right) = 0$.

Tables

Table D-1: Distribution of behavior regarding the behavioral measure of concern for morality

Number of lies	Reported number for state:						Frequency	Share	Classification
	1	**2**	**3**	**4**	**5**	**6**			
0 (Honest)	1	2	3	4	5	6	161	0.671	High-theta
1	6	2	3	4	5	6	6	0.038	Low-theta
	2	2	3	4	5	6	3		
2	6	6	3	4	5	6	12	0.104	Low-theta
	1	2	3	6	6	6	2		
	1	3	3	5	5	6	1		
	1	4	4	4	5	6	1		
	5	6	3	4	5	6	2		
	6	5	3	4	5	6	1		
	3	2	3	5	5	6	1		
	3	3	3	4	5	6	5		
3	6	6	6	4	5	6	11	0.050	Low-theta
	4	2	3	6	6	6	1		
4	6	6	6	6	5	6	3	0.017	Low-theta
	6	5	5	5	5	6	1		
5	6	6	6	6	6	6	15	0.067	Low-theta
	2	3	4	5	6	6	1		
Lied in a self-harmful manner	1	2	3	4	3	3	1	0.054	High-theta
	1	2	3	4	4	4	1		
	1	2	3	4	5	5	1		
	1	3	2	5	4	6	1		
	1	4	2	4	5	6	1		
	1	4	6	3	5	6	1		
	2	1	3	4	5	6	1		
	3	4	5	4	6	2	1		
	5	1	3	6	4	2	1		
	5	2	3	4	1	6	1		
	5	4	6	4	6	5	1		
	6	2	5	5	1	3	1		
	6	6	6	6	6	5	1		

Table D-2: Relationship between participation decision/reservation wage and θ^{Exp} (Hurdle model)

Dependent variable:	Participate	Reservation wage	Participate	Reservation wage	Reservation wage	Reservation wage
	(1)	(2)	(3)	(4)	(5)	(6)
Low-theta (θ_L^{Exp})	1.024***	-0.494*	0.925***	-0.362	-0.018	-0.060
	(4.64)	(-1.79)	(4.64)	(-0.99)	(-0.22)	(-0.39)
Period			-0.019**	-0.047		-0.050***
			(-2.40)	(-1.64)		(-5.74)
Period * θ_L^{Exp}			0.012	-0.015		0.005
			(1.14)	(-0.40)		(0.54)
Constant	0.295**	4.056***	0.449***	4.425***	2.909***	3.312***
	(2.41)	(20.20)	(3.86)	(14.76)	(43.79)	(25.63)
Sigma		2.64***		2.630***	0.609***	0.571***
		(7.72)		(7.69)	(10.57)	(10.58)
Market	Immoral	Immoral	Immoral	Immoral	Neutral	Neutral
N	2520	1755	2520	1755	1077	1077
LL (pseudo)	-1427.9	-4194.1	-1424.0	-4187.4	-993.9	-924.3
p-value: t + t* θ_L^{Exp}= 0			0.422	0.001		0.0000

*Notes: Estimates from Craggs Model: (1) and (3) are probit models; (2), (4), (5) and (6) are truncated linear regressions (truncated from above at 50 CHF). Models (1) to (4) use only data from the immoral markets, models (5) and (6) use only data from the neutral markets. For neutral markets, we do not report the regression of market participation as we have only 3 incidences where a subject did not participate. Independent variables: Low-theta in {0, 1}, Period between 1 and 15. Standard errors clustered at market level; z-statistics in parentheses; * - p < 0.1; ** - p < 0.05; *** - p < 0.01.*

Table D-3: Relationship between the behavioral measure of concern for morality and outcomes in the experimental labor markets

Dependent variable:	Employment rate			
	(1)	(2)	(3)	(4)
Number of lies				
1 lie	0.201	0.128	0.004	-0.008
	(1.05)	(0.53)	(0.11)	(-0.12)
2 lies	0.220**	0.182	-0.011	-0.049
	(2.17)	(1.39)	(-0.21)	(-0.84)
3 lies	0.398***	0.351***	-0.296***	-0.326***
	(5.29)	(3.57)	(-12.89)	(-15.52)
4 lies	0.392***	0.279**	0.171***	0.167***
	(4.94)	(2.61)	(7.43)	(7.18)
5 lies	0.286***	0.233**	0.037	0.018
	(3.26)	(2.48)	(0.49)	(0.17)
self-harmful lies	0.252**	0.182	-0.0516	-0.001
	(2.53)	(1.37)	(-1.25)	(-0.04)
Constant	0.475***	0.517***	0.829***	0.806***
	(11.6)	(32.97)	(36.11)	(35.35)
Market	Immoral	Immoral	Neutral	Neutral
N	168	168	72	72
R^2	0.121	0.263	0.085	0.250
Market FE	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Models (1) and (2) use only data from the immoral markets, models (3) and (4) use only data from the neutral markets. Standard errors clustered at market level; t-statistics in parentheses; * - $p < 0.1$; ** - $p < 0.05$; *** - $p < 0.01$.*

Table D-4: Relationship between θ^{Exp} and welfare

Dependent variable:	Market welfare		Self-reported happiness	
	(1)	(2)	(3)	(4)
Low-theta (θ_L^{Exp})	-1.07**	-0.99*	0.64	0.43
	(-2.42)	(-1.89)	(1.03)	(0.60)
Immoral market (Im)	8.83***		-0.07	
	(9.93)		(-0.25)	
θ_L^{Exp} * Im	8.12***	9.92***	-0.48	-0.20
	(3.22)	(3.57)	(-0.64)	(-0.22)
N	240	240	240	240
R^2	0.134	0.235	0.006	0.113
p-value: θ_L^{Exp} $+\theta_L^{Exp}$*IM=0	0.007	0.002	0.719	0.670
Market FE	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Independent variables: Low-theta in {0, 1}, Immoral market in {0, 1}. Standard errors clustered at market level; t-statistics in parentheses; * - p < 0.1; ** - p < 0.05; *** - p < 0.01.*

Table D-5: Relationship between θ^{Exp} and market income/welfare, depending on the behavior of other market participants

Dependent variable:	Market income	Market welfare
Low-theta (θ_L^{Exp})	-3.679	-1.455
	(-1.05)	(-0.39)
More θ_H types	4.334*	0.480
	(2.05)	(0.27)
θ_L^{Exp} * More θ_H types	16.348***	13.270***
	(3.53)	(2.85)
Constant	18.046***	13.993***
	(11.61)	(10.09)
N	168	168
R^2	0.089	0.072
p-value: More θ_H types + θ_L^{Exp} * More θ_H types=0	0.0002	0.0068

*Notes: Coefficient estimates of linear regression models. Low-theta in {0, 1}, More ϑ_H types in {0, 1}. Standard errors clustered at market level; t-statistics in parentheses; * - p < 0.1; ** - p < 0.05; *** - p < 0.01.*

Table D-6: Description and summary statistics of survey scales

Variable	Number of items	Mean (Sd)	Interpretation
Protected value 1	5	0.75 (0.19)	First measure of protected value, = 1 if the person finds that the behavior of a banker who recommends sub-optimal assets to his clients because he has larger margins on them is: very outrageous, very blameworthy, very immoral, not at all acceptable and not at all praiseworthy.
Protected value 2	4	0.63 (0.18)	Second measure of protected value, = 1 if the person thinks that truthfulness is a value that cannot be sacrificed.
Work ethics 1	1	0.38 (0.29)	= 1 if the person thinks that people are generally dishonest.
Work ethics 2	1	0.55 (0.36)	= 1 if the person thinks that calling sick to have a free day at work is really bad.
HEXACO sincerity	3	0.59 (0.19)	= 1 if the person is very sincere.
HEXACO fairness	3	0.68 (0.22)	= 1 if the person is very fair.
HEXACO greed avoidance	2	0.58 (0.23)	= 1 it the person is not at all greedy.
HEXACO modesty	2	0.66 (0.22)	= 1 if the person is very modest.
Charity attitude index	9	0.69 (0.13)	= 1 if the person's attitude towards charities is very positive.

Source: on-line survey. Notes: N = 237, every survey scale is obtained by averaging the individuals' answers to every item constituting the scale under consideration.

Table D-7: Relationship between θ^{Sur} and outcomes in the experimental labor market

Dependent variable:	**Employment rate**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Type survey (θ^{Sur})	-0.008	0.036	0.104	0.150	0.0103	0.015
	(-0.06)	(0.22)	(0.45)	(0.75)	(0.06)	(0.06)
Immoral market (Im)	-0.046		0.013		0.133	
	(-0.37)		(0.07)		(0.59)	
θ^{Sur} * Im	-0.431	-0.366	-0.558	-0.537	-0.648*	-0.505
	(-1.64)	(-1.14)	(-1.53)	(-1.26)	(-1.69)	(-1.05)
Aggregation θ^{Sur}	Factor Analysis	Factor Analysis	Equal weight	Equal weight	Theta-Exp	Theta-Exp
N	237	237	237	237	237	237
R^2	0.137	0.294	0.132	0.291	0.133	0.293
p-value: θ^{Sur} + θ^{Sur}*Im = 0	0.0569	0.237	0.0924	0.311	0.0708	0.252
Market FE	No	Yes	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Dependent variables: Models differ in how θ^{Sur} is constructed from the nine psychological measures. Columns (1) and (2) report our main results, using factor analysis to aggregate the psychological measures. Columns (3) and (4) give the result if equal weight is given to each measure instead. Columns (5) and (6) give the results if weights are determined by a regression of the survey measures on θ^{Exp}. Immoral market in {0, 1}. θ^{Sur} is in [0,1], where 0 means immoral and 1 means moral. Standard errors clustered at market level; t-statistics in parentheses; * - p < 0.1; ** - p < 0.05; *** - p < 0.01.*

Table D-8: Relationship between participation decision/reservation wage and θ^{Sur} (Hurdle model) in the immoral market

Dependent variable:	Participate	Reservation wage	Participate	Reservation wage	Participate	Reservation wage
	(1)	(2)	(3)	(4)	(5)	(6)
Type survey (θ^{Sur})	-1.614**	0.432	-1.971**	0.571	-2.177**	0.373
	(-2.44)	(0.60)	(-2.08)	(0.61)	(-2.13)	(0.33)
Constant	1.326***	3.674***	1.806***	3.515***	1.841***	3.657***
	(3.98)	(10.39)	(2.99)	(6.12)	(3.01)	(5.25)
Sigma		2.679***		2.679***		2.680***
		(7.71)		(7.72)		(7.70)
Aggregation θ^{Sur}	Factor Analysis	Factor Analysis	Equal weight	Equal weight	Theta-Exp	Theta-Exp
N	2475	1711	2475	1711	2475	1711
LL (pseudo)	-1478.3	-4114.1	-1488.2	-4114.2	-1490.9	-4114.6

*Notes: Estimates from Craggs Model: Regressions (1), (3) and (5) are probit models, regressions (2), (4) and (6) are truncated linear regressions (truncated from above at 50 CHF). Independent variables: Regressions differ in how θ^{Sur} is constructed from the nine psychological measures. Columns (1) and (2) report our main results, using factor analysis to aggregate the psychological measures. Columns (3) and (4) give the result if equal weight is given to each measure instead. Columns (5) and (6) give the results if weights are determined by a regression of the survey measures on θ^{Exp}.. θ^{Sur} is in [0,1], where 0 means immoral and 1 means moral. Standard errors clustered at market level; z-statistics in parentheses; * - $p < 0.1$; ** - $p < 0.05$; *** - $p < 0.01$. In the moral market, the coefficient of θ^{Sur} is not significant for any of the above specifications.*

Table D-9: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types, robustness checks aggregation Θ^{Sur}

Dependent variable:	**Willingness to work for industry *j***				**Willingness to work for firm *j***			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Perceived	0.089	0.090	0.122	0.122	0.396***	0.391***	0.337***	0.327***
immorality (I(j))	(1.05)	(1.05)	(1.37)	(1.38)	(4.89)	(4.87)	(4.07)	(3.96)
Type survey	-0.154*	-0.149*	-0.131	-0.158	-0.200*	-0.267**	-0.243**	-0.298**
(Θ^{Sur})	(-1.95)	(-1.77)	(-1.46)	(-1.65)	(-1.81)	(-2.57)	(-2.12)	(-2.58)
$\Theta^{Sur} * I(j)$	-0.583***	-0.583***	-0.671***	-0.671***	-0.998***	-0.990***	-0.961***	-0.944***
	(-4.57)	(-4.57)	(-4.78)	(-4.78)	(-8.08)	(-8.06)	(-7.23)	(-7.10)
Aggregation Θ^{Sur}	Equal weight	Equal weight	Theta-Exp	Theta-Exp	Equal weight	Equal weight	Theta-Exp	Theta-Exp
N	4715	4715	4715	4715	5064	5064	5064	5064
R^2	0.098	0.119	0.105	0.127	0.126	0.158	0.121	0.153
Control variables	No	Yes	No	Yes	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Independent variables: Column (1), (2), (5) and (6) give the result if equal weight is given to each measure. Observations where subjects did not know the firm ("I don't know this organization") or did not fill out the questionnaire are excluded. Column (3), (4), (7) and (8) give the results if weights are determined by a regression of the survey measures on θ_L^{Exp}. Θ^{Sur} is in [0,1] where 0 means immoral and 1 means moral, willingness to work is in {0, 0.25, 0.5, 0.75, 1} where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study. Standard errors clustered at individual level; t-statistics in parentheses; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table D-10: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types, robustness checks classification firms immorality

Dependent variable:	**Willingness to work for firm *j***			
	(1)	**(2)**	**(3)**	**(4)**
Perceived immorality ($I_{Alt}(j)$)	-0.157***	-0.156***	0.141**	0.138**
	(-4.18)	(-4.20)	(2.73)	(2.68)
Type experiment (Θ^{Exp})	-0.045	-0.053**		
	(-1.62)	(-1.99)		
Θ^{Exp} * $I_{Alt}(j)$	-0.174***	-0.175***		
	(-3.99)	(-4.04)		
Type survey (Θ^{Sur})			-0.184**	-0.222***
			(-2.35)	(-2.90)
Θ^{Sur} * $I_{Alt}(j)$			-0.848***	-0.840***
			(-8.52)	(-8.45)
N	5064	5064	5064	5064
R^2	0.107	0.138	0.125	0.157
Control variables	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Perceived immorality is calculated different then in our main analysis: clients that did not know a firm are classified as giving a neutral rating. Dependent variable: Willingness to work is in {0, 0.25, 0.5, 0.75, 1} where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm ("I don't know this organization") or did not fill out the questionnaire are excluded. Independent variables: (1) and (2) use Θ^{Exp} to classify participants, where Θ^{Exp}=0 for low- theta types and Θ^{Exp}=1 for high-theta types, while (3) and (4) use Θ^{Sur} (in [0,1]) instead. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study. Standard errors clustered at individual level; t-statistics in parentheses; * p < 0.1; ** p < 0.05; *** p < 0.01.*

Table D-11: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types, robustness checks "I don't know this organization"

Dependent variable:	**Willingness to work for firm *j***							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Perceived	-0.123***	-0.123***	0.088**	0.088**	-0.092*	-0.092*	0.159**	0.159**
immorality (I(j))	(-4.51)	(-4.51)	(2.35)	(2.35)	(-1.92)	(-1.91)	(2.19)	(2.18)
Type experiment	-0.034	-0.040*			-0.039	-0.056		
(Θ^{Exp})	(-1.51)	(-1.87)			(-0.56)	(-0.86)		
$\Theta^{Exp} * I(j)$	-0.131***	-0.131***			-0.158**	-0.158**		
	(-4.08)	(-4.07)			(-2.61)	(-2.59)		
Type survey			-0.127**	-0.165***			-0.156	-0.204
(Θ^{Sur})			(-2.02)	(-2.63)			(-1.07)	(-1.10)
$\Theta^{Sur} * I(j)$			-0.613***	-0.613***			-0.792***	-0.792***
			(-8.29)	(-8.28)			(-4.95)	(-4.92)
N	6162	6162	6162	6162	1352	1352	1352	1352
R^2	0.096	0.123	0.110	0.139	0.086	0.157	0.110	0.182
Control variables	No	Yes	No	Yes	No	Yes	No	Yes

*Notes: Coefficient estimates of linear regression models. Dependent variable: Willingness to work is in {0, 0.25, 0.5, 0.75, 1} where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not fill out the questionnaire are excluded. Columns (1) to (4): Observations where subjects did not know the firm ("I don't know this organization") are classified as having willingness to work of 0.5. Columns (5) to (8): only participants that did know all firms (N=52) are included. Independent variables: (1), (2), (5) and (6) use Θ^{Exp} to classify participants, where Θ^{Exp}=0 for low- theta types and Θ^{Exp}=1 for high-theta types, while (3), (4), (7) and (8) use Θ^{Sur} (in [0,1]) instead. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study. Standard errors clustered at individual level; t-statistics in parentheses; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

Table D-12: Summary of main variables from the Swiss Labor Force Survey, overall and by industry

Industries	N	Real gross hourly wage (in 2010 CHF)	Ln of the real gross hourly wage	Age	Male	Married	Edu high	Edu middle	Swiss	Tenure	Full-time equivalent	Board member	Perceived immorality I(j)
All industries	18514	41.35	3.6	42.37	0.57	0.58	0.36	0.5	0.63	9.62	0.86	0.16	-0.18
		(20.24)	(0.56)	(11.25)	(0.49)	(0.49)	(0.48)	(0.5)	(0.48)	(9.26)	(0.24)	(0.37)	(0.22)
Processing of tea and coffee	75	40.73	3.65	39.13	0.78	0.63	0.29	0.54	0.45	9.42	0.97	0.07	-0.11
		(13.65)	(0.34)	(9.18)	(0.42)	(0.49)	(0.46)	(0.5)	(0.5)	(9.93)	(0.12)	(0.25)	(0.41)
Manufacture of tobacco products	88	52.94	3.91	38.95	0.61	0.5	0.54	0.32	0.48	9.44	0.93	0.03	0.47
		(17.19)	(0.38)	(7.14)	(0.49)	(0.5)	(0.5)	(0.47)	(0.5)	(7.88)	(0.09)	(0.16)	(0.42)
Manufacture of paper and paperboard	54	38.83	3.63	43.73	0.83	0.72	0.24	0.61	0.48	15.48	0.97	0.07	-0.06
		(10.25)	(0.24)	(11.07)	(0.38)	(0.45)	(0.43)	(0.49)	(0.5)	(12.65)	(0.08)	(0.27)	(0.37)
Manufacture of weapons and ammunitions	57	48.91	3.83	46.61	0.96	0.64	0.53	0.43	0.88	15.94	0.97	0.08	0.71
		(18.69)	(0.34)	(10.55)	(0.21)	(0.48)	(0.5)	(0.5)	(0.32)	(13.46)	(0.04)	(0.28)	(0.4)
Manufacture of electronic components	656	45	3.73	43.49	0.69	0.64	0.51	0.37	0.56	10.77	0.92	0.05	-0.01
		(18.62)	(0.42)	(10.73)	(0.46)	(0.48)	(0.5)	(0.48)	(0.5)	(9.59)	(0.15)	(0.21)	(0.42)
Construction of buildings	2,022	38	3.58	43.63	0.92	0.7	0.22	0.47	0.48	11.14	0.95	0.14	-0.28
		(12.91)	(0.45)	(10.74)	(0.28)	(0.46)	(0.41)	(0.5)	(0.5)	(9.76)	(0.15)	(0.34)	(0.39)
Maintenance and repair of motor vehicles	1,508	35.11	3.49	41.02	0.82	0.59	0.21	0.71	0.66	11.28	0.92	0.23	-0.28
		(13.03)	(0.37)	(12.06)	(0.38)	(0.49)	(0.4)	(0.45)	(0.47)	(10.27)	(0.21)	(0.42)	(0.4)
Wholesale of tobacco products	51	62.67	4.02	43.14	0.57	0.72	0.53	0.47	0.62	11.64	0.94	0.12	0.44
		(29.55)	(0.52)	(9.5)	(0.5)	(0.45)	(0.5)	(0.5)	(0.49)	(8.12)	(0.14)	(0.33)	(0.38)
Wholesale of clothing and footwear	166	39.62	3.56	40.09	0.37	0.46	0.41	0.52	0.51	7.28	0.83	0.16	0.1
		(19.26)	(0.49)	(10.99)	(0.48)	(0.5)	(0.49)	(0.5)	(0.5)	(6.45)	(0.26)	(0.37)	(0.46)
Wholesale of perfume and cosmetics	200	51.02	3.78	41.3	0.26	0.61	0.54	0.4	0.44	8.25	0.88	0.13	0.12
		(27.74)	(0.57)	(10.64)	(0.44)	(0.49)	(0.5)	(0.49)	(0.5)	(7.18)	(0.19)	(0.34)	(0.37)

See next page for the rest of the table.

Industries	N	Real gross hourly wage (in 2010 CHF)	Ln of the real gross hourly wage	Age	Male	Married	Edu high	Edu middle	Swiss	Tenure	Full-time equivalent	Board member	Perceived immorality I(j)
Wholesale of watches and jewelry	91	44 (19.12)	3.7 (0.4)	44.21 (10.69)	0.39 (0.49)	0.58 (0.5)	0.45 (0.5)	0.47 (0.5)	0.47 (0.5)	8.25 (7.96)	0.87 (0.19)	0.16 (0.37)	0.04 (0.41)
Hotels and similar accommodation	1,514	27 (10.51)	3.22 (0.47)	40.47 (11.45)	0.39 (0.49)	0.55 (0.5)	0.18 (0.38)	0.57 (0.5)	0.4 (0.49)	6.94 (7.56)	0.84 (0.27)	0.16 (0.36)	-0.34 (0.37)
Restaurants and mobile food activities	3,137	25 (10.37)	3.14 (0.48)	40.67 (11.96)	0.43 (0.5)	0.56 (0.5)	0.15 (0.36)	0.56 (0.5)	0.46 (0.5)	6.77 (7.6)	0.8 (0.29)	0.21 (0.4)	-0.33 (0.37)
Monetary intermediations	4,146	56.56 (22.02)	3.95 (0.46)	42.53 (10.49)	0.59 (0.49)	0.56 (0.5)	0.56 (0.5)	0.42 (0.49)	0.72 (0.45)	10.87 (9.7)	0.91 (0.19)	0.19 (0.39)	0.11 (0.4)
Credit granting	44	52 (22.91)	3.86 (0.41)	41.15 (8.81)	0.41 (0.5)	0.53 (0.5)	0.46 (0.5)	0.51 (0.51)	0.68 (0.47)	6.02 (5.51)	0.86 (0.2)	0.05 (0.23)	0.15 (0.41)
Non-life insurance	1,484	50 (19.05)	3.84 (0.4)	42.11 (10.97)	0.49 (0.5)	0.54 (0.5)	0.53 (0.5)	0.44 (0.5)	0.78 (0.41)	9.91 (8.97)	0.88 (0.2)	0.08 (0.27)	-0.13 (0.44)
General public administration activities	2,615	44.24 (17.3)	3.69 (0.57)	45.16 (10.99)	0.48 (0.5)	0.59 (0.49)	0.43 (0.5)	0.51 (0.5)	0.91 (0.29)	10.62 (9.58)	0.79 (0.28)	0.13 (0.34)	-0.41 (0.36)
Gambling and betting activities	91	40.57 (18.55)	3.61 (0.45)	42.58 (10.26)	0.6 (0.49)	0.51 (0.5)	0.29 (0.46)	0.55 (0.5)	0.46 (0.5)	7.07 (5.4)	0.86 (0.22)	0.09 (0.29)	0.42 (0.41)
Organization and operation of sport facilities for indoor and outdoor sports events	288	32.76 (14.53)	3.36 (0.67)	43.48 (13.19)	0.48 (0.5)	0.52 (0.5)	0.27 (0.44)	0.58 (0.49)	0.68 (0.47)	7.61 (8.55)	0.7 (0.35)	0.08 (0.27)	-0.49 (0.4)
Fitness facilities	227	31.15 (16.31)	3.29 (0.58)	40.2 (11.21)	0.25 (0.43)	0.6 (0.49)	0.33 (0.47)	0.57 (0.5)	0.64 (0.48)	5.69 (6.53)	0.5 (0.35)	0.19 (0.39)	-0.35 (0.38)

Source: Weighed data from the SLFS, years: 2010, 2012, 2014 and 2016 (wage) and our own survey (Perceived immorality). Same sample used in Table 15, Figures 13 and D1. Notes: N=number of observations per industry in the SLFS dataset, Male in {0, 1}, Married in {0, 1}, Education high: higher vocational education and training or university/college, Education middle: apprenticeship, full-time vocational school, matura or pedagogical training, Education low (reference category): compulsory schooling or pre-vocational education, Swiss in {0, 1}, Tenure = number of years in the firm, Full-time equivalent = (working hours /42), set to 1 for working hours >= 42, Board member in {0, 1}, Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Standard deviations in parentheses

Table D-13: Perceived immorality of firms

Firms	Perceived immorality I(j)	Firms	Perceived immorality I(j)
Marlboro	0.54	**Swisscom**	-0.07
Monsanto	0.52	**Firmenich**	-0.09
Glencore	0.46	**Winterthur Assurance**	-0.1
Philip Morris	0.46	**Swiss Life**	-0.13
Nestlé	0.39	**Swatch**	-0.17
Tamoil	0.37	**Adecco**	-0.18
Syngenta	0.23	**ABB**	-0.2
UBS	0.19	**Migros**	-0.38
Novartis	0.18	**WWF**	-0.66
Credit Suisse	0.17	**Pro Juventute**	-0.66
Roche	0.13	**Pro Natura**	-0.67
Holcim	0.03	**UNICEF**	-0.72
Ernst and Young	-0.05	**Red cross**	-0.81

Source: own survey. Notes: Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral.

Table D-14: Ordered logit estimates of willingness to work for diverse industries and firms on perceived immorality and moral types

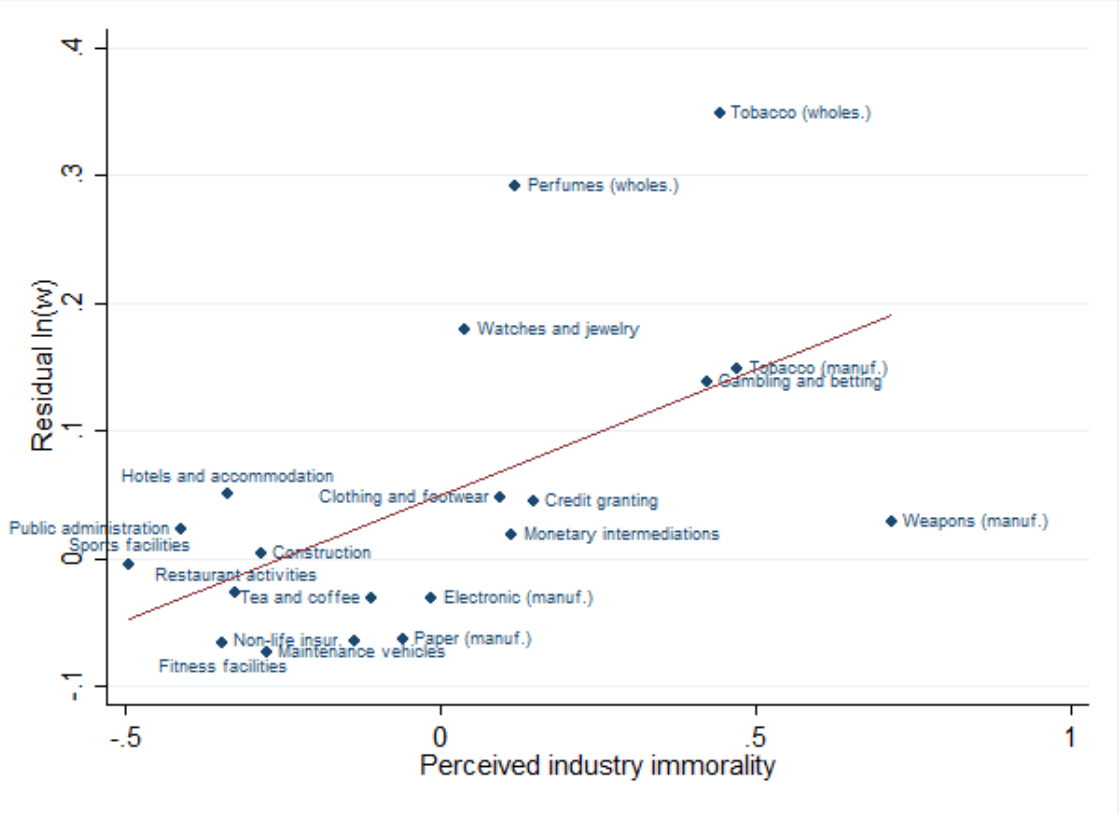
Dependent variable:	Willingness to work for industry *j*				Willingness to work for firm *j*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Perceived immorality (I(j))	-1.474***	-1.502***	-1.403***	-1.434***	-0.799***	-0.797***	-0.89***	-0.893***
	(6.88)	(6.97)	(9.70)	(9.81)	(3.92)	(3.88)	(6.54)	(6.50)
Type experiment (Θ^{Exp})	-0.186	-0.188			-0.259	-0.3*		
	(1.4)	(1.45)			(1.59)	(1.9)		
$\Theta^{Exp} * I(j)$	-0.608**	-0.613**			-0.902***	-0.918***		
	(2.46)	(2.46)			(3.77)	(3.79)		
Type survey (Θ^{Sur})			-0.323***	-0.35***			-0.273**	-0.316**
			(2.67)	(2.93)			(1.98)	(2.26)
$\Theta^{Sur} * I(j)$			-1.085***	-1.089***			-1.219***	-1.232***
			(5.24)	(5.23)			(5.97)	(6.00)
N	4,715	4,715	4,715	4,715	5,064	5,064	5,064	5,064
Control variables	No	Yes	No	Yes	No	Yes	No	Yes

*Notes: Dependent variable: Willingness to work is in {0, 0.25, 0.5, 0.75, 1} where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm ("I don't know this organization") are dropped. Independent variables: (1), (2), (5) and (6) use Θ^{Exp} to classify participants, where Θ^{Exp}=0 for low- theta types and Θ^{Exp}=1 for high-theta types, while (3), (4), (7) and (8) use Θ^{Sur} (in [0,1]) instead. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study. Standard errors clustered at individual level; z-statistics in parentheses; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

APPENDIX D

Figures

Figure D-1: Correlation between wages (residual) and perceived industry immorality



Source: Weighted data from the SLFS, years: 2010, 2012, 2014 and 2016 (wage) and our own survey (perceived immorality). Notes: Cross-section with N = 18514. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Dependent variable: logarithm of real gross hourly wage (in 2010 CHF). Control variables: Male in {0, 1}, Married in {0, 1}, Edu high in {0, 1}: Higher vocational education and training, or University college, Edu middle in {0, 1}: Graduate school, apprenticeship, full-time vocational school, Matura, Teacher seminar, Edu low (reference category): compulsory schooling or pre-vocational education, Swiss in {0, 1}, Tenure = number of years in the firm, Full-time equivalent = working hours/42, set to 1 for working hours >= 42, Board member in {0, 1}, seven industrial branch dummies (manufacturing; wholesale and retail trade; repair of motor vehicle and motorcycles; financial and insurance activities; accommodation and food service activities; construction; public administration, defense and compulsory social security; arts, entertainment and recreation). Number of observations: Credit granting=44, Monetary intermediations=4146, Gambling and betting activities=91, Manufacture of tobacco products=88, Manufacture of weapons and ammunitions=57, Wholesale of tobacco products=51, Non-life insurance=1484, Organization and operation of sport facilities=288, Processing of tea and coffee=75, Manufacture of electronic components=656, Wholesale of perfume and cosmetics=200, Wholesale of clothing and footwear=166, Wholesale of watches and jewelry =91, Manufacture of paper and paperboard=54, Maintenance and repair of motor vehicles=1508, Construction of buildings=2022, Restaurant and mobile food activities=3137, Hotels and similar accommodation=1514, Fitness facilities=227, General public administration activities=2615. For credit granting: extrapolation based on less than 50 observations; this result must be interpreted with great caution.

Figure D-2: Labor supply for neutral and immoral work in the laboratory (last 5 periods only)

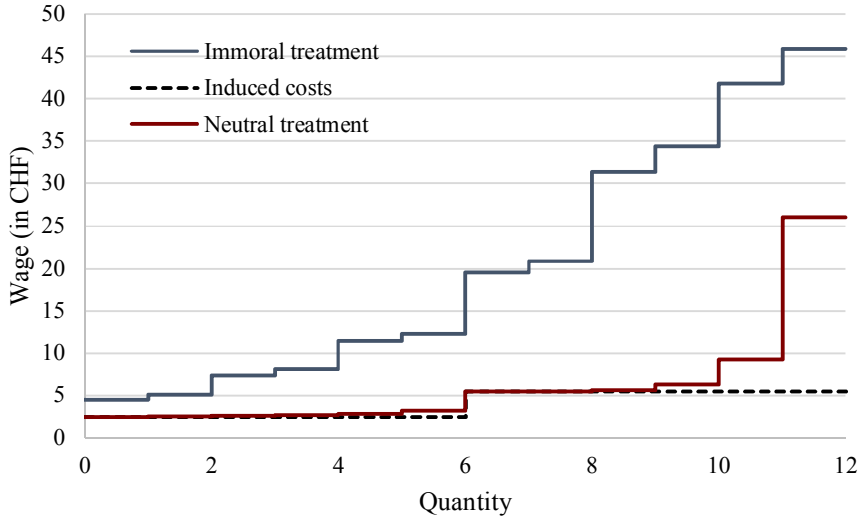


Figure D-3: Employment rate by the two types in the neutral treatment

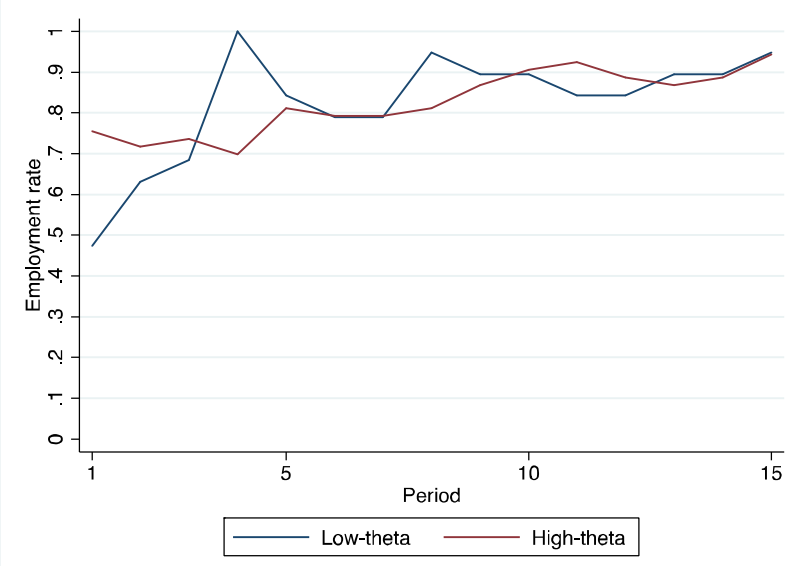


Figure D-4: Probability distribution of θ^{Sur}

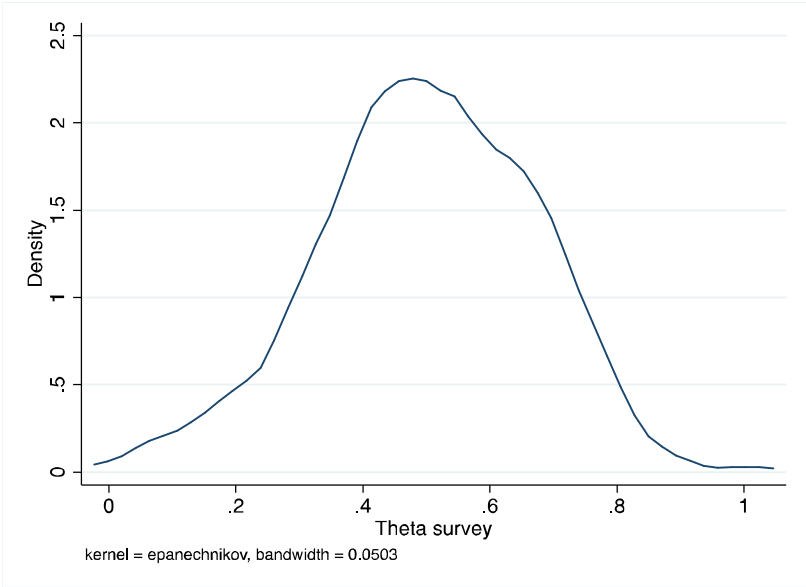
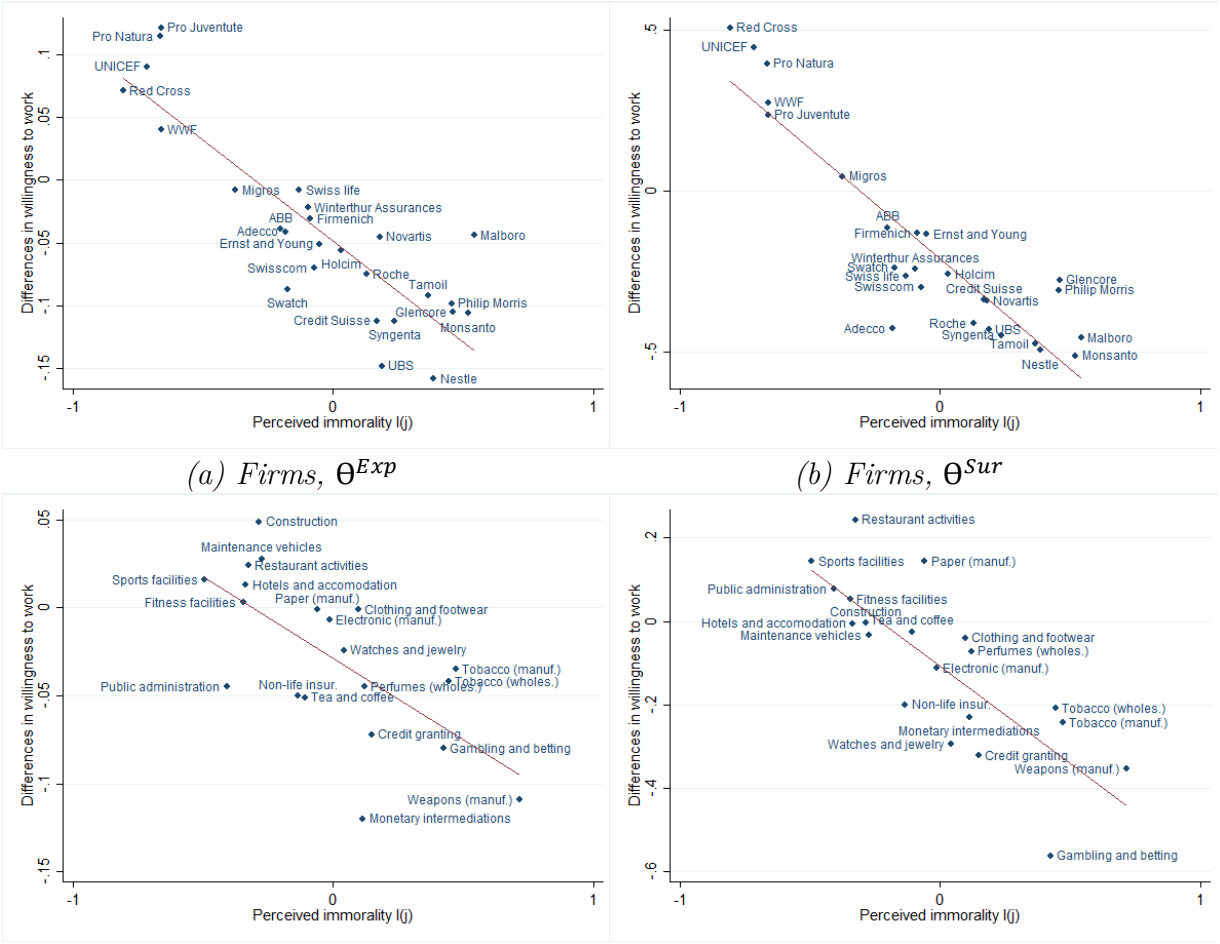


Figure D-5. Difference in willingness to work for diverse firms and industries between moral and immoral types, with respect to the perceived immorality of these institutions, and controlling for demographics.



(a) Firms, Θ^{Exp} *(b) Firms,* Θ^{Sur}

(c) Industries, Θ^{Exp} *(d) Industries,* Θ^{Sur}

Source: Survey study (Perceived immorality), on-line survey (Willingness to work, Θ^{Sur}*), Laboratory experiment (*Θ^{Exp}*). Notes: Differences in willingness to work: Coefficient estimates of linear regression models of the participants' willingness to work for different industries (a and b) or firms (b and c) on* Θ_H^{Exp} *(a and c) or* Θ_H^{Sur} *(b and d), and control variables (age, gender, Swiss nationality, subject of study). Dependent variable: Willingness to work is in {0, 0.25, 0.5, 0.75, 1} where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm ("I don't know this organization") or did not fill out the questionnaire are excluded. Independent variables: a and c use* Θ^{Exp} *to classify participants, where* Θ^{Exp}*=0 for low- theta types and* Θ^{Exp}*=1 for high-theta types, while b and d use* Θ^{Sur} *in [0,1] instead. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Standard errors clustered at individual level.*

References

Abeliotis, K., Lasaridi, K. and Chroni, C., (2014) "Attitudes and behavior of Greek households regarding food waste prevention," Waste Management and Research, 32: 237-240.

Abrantes-Metz, R. M., Kraten, M., Metz, A. D. and Seow, G. S. (2012) "Libor manipulation?," Journal of Banking and Finance, 36: 136–150.

Ambühl, S. (2017) "An Offer You Can't Refuse? Incentives Change What We Believe," Working Paper.

Andreoni, J., Nikiforakis, N. and Stoop, J. (2017) "Are the Rich More Selfish than the Poor, or Do They Just Have More Money? A Natural Field Experiment," NBER Working Paper No. 23229.

Ariely, D., Bracha, A. and Meier, S. (2009) "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially," American Economic Review, 99(1): 544-555.

Armantier, O. and Treich, N. (2013) "Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging," European Economic Review, 62:17-40.

Ashton, M.C. and Lee, K. (2009) "The HEXACO-60: A Short Measure of the Major Dimensions of Personality," Journal of Personality Assessment, 91(4): 340-345.

Auten, G., Sieg, H., and Clotfelter, C. (2002) "Charitable giving, income, and taxes: an analysis of panel data," American Economic Review, 92(1): 371–382.

Bartling, B., Engl, F. and Weber, R. A. (2015) "Game form misconceptions are not necessary for a willingness-to-pay vs. willingness-to-accept gap," Journal of the Economic Science Association, 1(1): 72-85.

Bartling, B., Fehr, E. and Herz, H. (2014) "The intrinsic value of decision rights," Econometrica, 82(6), 2005-2039.

Bartling, B., Valero, V., and Weber, R. A. (2018) "Income Effects on Socially Responsible Market Behavior," working paper.

Bartling, B., Weber, R. A. and Yao, L. (2015) "Do Markets Erode Social Responsibility?," Quarterly Journal of Economics, 130(1): 219-66.

Bates, Cl. and Rowell, A. (1998) "Tobacco explained: The truth about the tobacco industry ... in its own words." Scotland: ASH.

Becker, G. M., DeGroot M. H., and J. Marschak (1964) "Measuring Utility by a Single-Response Sequential Method," Behavioral Sciences, 9 (3): 226-32.

Bell, D. E. (1982) "Regret in decision making under uncertainty," Operations Research, 30, 961-981.

Bell, D. E. (1985) "Disappointment in decision making under uncertainty. Operations Research, 33(1).

Benedict, M.E., McClough, D and McClough, A.C. (2006) "The price of morals: An empirical investigation of industry sectors and perceptions of moral satisfaction – do business economists pay for morally sytisfying employment?," The American Economist, 50(1): 21-36.

Besley, T. and Ghatak, M. (2005) "Competition and Incentives with Motivated Agents," American Economic Review, 95(3): 616-636.

Bock, O., Baetge, I., and Nicklisch, A. (2014) "Hroot: Hamburg registration and organization on-line tool," European Economic Review, 71: 117-120.

Botti, S. (2004) "The psychological pleasure and pain of choosing: when people prefer choosing at the cost of subsequent outcome satisfaction. Journal of personality and social psychology, 87(3), 312.

Bradley M.M. and Lang, P.J. (1994) "Measuring emotion: the self-assessment manikin and the semantic differential," Journal of Behavioral Therapy and Experimental Psychiatry, 25(1): 49-59.

Brashear, G., Green, L., and Webb, J. (2000) "Development and validation of scales to measure attitudes influencing monetary donations to charitable organizations," Journal of the Academy of Marketing Science, 28(2): 299-309.

Brehm, J. W. (1956) "Postdecision changes in the desirability of alternatives," Journal of Abnormal and Social Psychology, 52(3), 384.

Brenner, L., Rottenstreich, Y. and Sood, S. (1999) "Comparison, grouping, and preference. Psychological Science, 10(3).

Brook Lyndhurst (2007) "Food behavior consumer research – Findings from the quantitative survey," Briefing Paper, WRAP, UK.

Brown, D. R. (1953) "Stimulus-similarity and the anchoring of subjective scales," The American Journal of Psychology, 66(2), 199-214.

Brun, F. (2018) "Exploring behaviors towards food waste," Doctoral Thesis.

Brun, F. and Weber, A. R. (2018) "Psychological depreciation and food waste," Working paper.

Carmon, Z., Wertenbroch, K., and Zeelenberg, M. (2003) "Option attachment: When deliberating makes choosing feels like losing," Journal of Consumer Research, 30(1), 15-29.

Carpenter, J. and Gong, E. (2016) "Motivating Agents: How Much Does the Mission Matter?," Journal of Labor Economics, 34(1): 211-236.

Cassar, L and Armouti-Hansen, J. (2018) "Optimal Contracting with Endogenous Project Mission," working Paper.

Cassar, L. (forthcoming) "Job Mission as a Substitute for Monetary Incentives: Benefits and Limits," Management Science.

Cohn, A., Fehr, E. and Maréchal, A. (2014) “Business culture and dishonesty in the banking industry,” Nature, 516: 86-89.

Colonnello, S., Curatola, G. and Gioffré, A. (2017) “Pricing Sin Stocks: Ethical Preference vs. Risk Aversion,” working paper.

Connolly, T., and Zeelenberg, M. (2002) “Regret in decision making,” Current Directions in Psychological Science, 11(6), 212-216.

De Hooge, I. E., Oostindjer, M., Aschemann-Witzel, J., Normann, A., Mueller Loose S., and Lengard Almli,V. (2017) "This apple is too ugly for me! Consumer preferences for suboptimal food products in the supermarket and at home," Food Quality and Preference, 56: 80-92.

Delfgaauw, J. and Dur, R. (2008) “Incentives and Workers’ Motivation in the Public Sector,” Economic Journal, 118(525): 171–191.

Département Fédéral de l’Intérieur [DFI] (2017) “Ordonnance du DFI sur l’hygiène dans les activités liées aux denrées alimentaires. ”

Dhar, R. (1997) “Consumer preference for a no-choice option,” Journal of Consumer Research, 24(2), 215-231.

Diehl, K., and Poynor, C. (2010) “Great expectations?! Assortment size, expectations, and satisfaction,” Journal of Marketing Research, 47(2), 312-322.

Dinnella, C., Torri, L., Caporale, G., and Monteleone, E. (2014) "An exploratory study of sensory attributes and consumer traits underlying liking for and perceptions of freshness for ready to eat mixed salad leaves in Italy," Food Research International, 59: 108-116.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011) „Individual risk attitudes: Measurement, determinants, and behavioral consequences,” Journal of the European Economic Association, 9(3): 522-550.

Edlund L. and Korn, E. (2002) “A Theory of Prostitution,” Journal of Political Economy, 110(1): 181-214.

European Commission (2011) “Roadmap to a resource efficient Europe.”

Evans D. (2011) “Blaming the consumer – once again: The social and material contexts of everyday food waste practices in some English households,” Critical Public Health, 21:429–540.

Evans D. (2012) “Beyond the throwaway society: Ordinary domestic practice and a sociological approach to household food waste,” Sociology, 46:41–56.

Falk A. and Szech, N. (2013) ‘‘Morals and Markets,’’ Science, 340: 707–711.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2017) “Global evidence on economic preferences,” The Quarterly Journal of Economics.

Fehrler, S. and Kosfeld, M. (2014) “Pro-Social Missions and Worker Motivation: An Experimental Study,” Journal of Economic Behavior and Organization, 100: 99-110.

Festinger, L. (1957) "A Theory of cognitive dissonance," Stanford, CA: Stanford University Press.

Fischbacher, U. (2007) "z-Tree: Zurich toolbox for ready-made economic experiments," Experimental Economics, 10(2), 171–178.

Food and Agriculture Organization of the United Nations [FAO] (2011), "Global food losses and food waste – Extent, causes and prevention." Rome.

Food and Agriculture Organization of the United Nations [FAO] (2013), "Food wastage footprint: Impacts on natural resources," Summary report.

Food Safety, Authority of Ireland (2009), "A research study into consumers' attitudes to food labelling."

Food Standards Agency [FSA] (2012), "Exploring food attitudes and behaviours in the UK: Findings from the food and you survey 2012." United Kingdom.

Frank, R. H. (1996) "What prices the moral high ground?," Southern Economic Journal, 63(1): 1-17.

Frank, R. H., Gilovich, T. and Regan, D. T. (1993) "Does Studying Economics Inhibit Cooperation?," Journal of Economic Perspectives, 7(2): 159-171.

Frey, B. S., and Stutzer, A. (2002) "What Can Economists Learn from Happiness Research? ," Journal of Economic Literature 40 (2): 402-435.

Garen, J. (1988) "Compensating Wage Differentials and the Endogeneity of Job Riskiness," Review of Economics and Statistics, 70(1): 9–16.

Garnett, T. (2011) "Where are the best opportunities for reducing greenhouse gas emissions in the food system (including the food chain)?," Food Policy, 36: 23-32.

Gibson, R., Tanner, C. and Wagner, A. (2013) "Preferences for truthfulness: Heterogeneity among and within individuals," American Economic Review, 103(1): 532-548.

Gilovich, T., and Medvec, V. H. (1995) "The experience of regret: What, when, and why," Psychological Review, 102(2), 379.

Gneezy, U. (2005) "Deception: the role of consequences," American Economic Review, 95(1): 384-394.

Gneezy, U., Rockenbach, B. and Serra-Garcia, M. (2013) "Measuring lying aversion," Journal of Economic Behavior and Organization, 93: 293-300.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M. and Toulmin, C. (2010) "Food security: The challenge of feeding 9 billion people," Science, New Series, 327(5967): 812-818.

Gooch, M., Felfel, A. and Marenick, N. (2010) „Food waste in Canada: Opportunities to increase the competitiveness of Canada's agri-food sector, while simultaneously improving the environment," Value Chain Management Centre, George Morris Center.

Goodman, J. K., Broniarczyk, S. M., Griffin, J. G., and McAlister, L. (2013) "Help or hinder? When recommendation signage expands consideration sets and heightens decision difficulty," Journal of Consumer Psychology, 23(2), 165-174.

Gosling, S.D., Rentfrow, P.J., and Swann Jr. W.B. (2003) "A very brief measure of the Big-Five personality domains," Journal of Research in Personality, 37: 504-528.

Graham-Rowe, E., Jessop, D. C., and Sparks, P. (2014) "Identifying motivations and barriers to minimizing household food waste," Resources, conservation and recycling, 84:15-23.

Graham-Rowe, E., Jessop, D. C., and Sparks, P. (2015) "Predicting household food waste reduction using an extended theory of planned behavior," Resources, Conservation and Recycling, 101: 194-202.

Greifeneder, R., Scheibehenne, B., and Kleber, N. (2010) "Less may be more when choosing is difficult: Choice complexity and too much choice," Acta psychologica, 133(1), 45-50.

Griffin, J. G., and Broniarczyk, S. M. (2010) "The slippery slope: The impact of feature alignability on search and satisfaction," Journal of Marketing Research, 47(2), 323-334.

Harcar, T. and Karakaya, F. (2005) "A cross-cultural exploration of attitudes toward product expiration dates," Psychology and Marketing, 22(4): 353–371.

Heath, D. (2016) "Contesting The Science of Smoking," The Atlantic, https://www.theatlantic.com/politics/archive/2016/05/low-tar-cigarettes/481116/

Hebrok, M., and Boks, C. (2017) "Household food waste: Drivers and potential intervention points for design–An extensive review," Journal of Cleaner Production, 151: 380-392.

Heidhues, P., Köszegi, B., and Murooka, T. (2016) "Inferior products and profitable deception," The Review of Economic Studies, 84(1), 323-356.

Hill, C. A. (2012) "Bankers behaving badly? The limits of regulatory reform," Review of Banking and Financial Law, 31: 675–691.

Hong, H. and Kacperczyk, M. (2009) "The Price of Sin: The Effects of Social Norms on Markets," Journal of Financial Economics, 93: 15-36.

Hsee, C. K., and Leclerc, F. (1998) "Will products look more attractive when presented separately or together? ," Journal of Consumer Research, 25(2), 175-186.

Huber, J., Payne, J. W., and Puto, C. (1982) "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," Journal of consumer research, 9(1), 90-98.

Hwang, H., Reed, W. R., and Hubbard, C. (1992) "Compensating Wage Differentials and Unobserved Productivity," Journal of Political Economy, 100(4): 835-858.

International Food Information Council Foundation [IFICF] (2007), "2007 Food and health survey, consumer attitudes towards food, nutrition and health, a trended survey."

Iyengar, S. S., and Lepper, M. R. (2000) “When choice is demotivating: Can one desire too much of a good thing? ,” Journal of Personality and Social Psychology, 79(6), 995.

Kahneman, D., and Krueger, A. B. (2006) “Developments in the Measurement of Subjective Well-Being,” Journal of Economic Perspectives 20 (1): 3-24.

Kahneman, D., Diener E., and Schwarz N. (Eds.) (1999) “Well-Being: Foundations of Hedonic Psychology,” New York: Russell Sage Foundation.

Kaiser, F. G., Ranney, M., Hartig, T., and Bowler, P. A. (1999) “Ecological behavior, environmental attitude, and feelings of responsibility for the environment,” European Psychologist, 4:59-74.

Kerley, D., Messer, K. D., Wansink, B., Kaiser, H. M. and Schulze, W. D. (2008) "Expiration dates and stigma: Why don't we observe hedonic markets for perishable products?," Conference paper (unpublished).

Kirchler, M., Huber, J., Stefan, M. and Sutter, M. (2016) “Market design and moral behavior,” Management Science, 62: 2615-2625.

Koivupuro, H. K., Hartikainen, H., Silvennoinen, K., Katajajuuri, J. M., Heikintalo, N., Reinikainen, A., and Jalkanen, L. (2012) “Influence of socio- d e mo g r a p h i c a behavioural and attitudinal factors on the amount of avoidable food waste generated in Finnish households,” International Journal of Consumer Studies, 36(2): 183-191.

Kole, A. P.W., Altintzoglou, T., Schelvis-Smit, R.A A. M. and Luten, J. B. (2009) "The effects of different types of product information on the consumer product evaluation for fresh cod in real life settings," Food Quality and Preference, 20: 187-194.

Kormos, C., and Gifford, R. (2014) “The validity of self-report measures of proenvironmental behavior: A meta-analytic review,” Journal of Environmental Psychology, 40, 359-371.

Krupka, E. and Weber, R. A. (2009) “The focusing and informational effects of norms on pro-social behavior,” Journal of Economic Psychology, 30(3): 307-320.

Leary, M.R., Diebels, K. J., and Jongman-Sereno, K. P. (2015) “Measures of concerns with public image and social evaluation”, in: Measures of Personality and Social Psychological Constructs, Elsevier Inc., 448-473. DOI: 10.1016/B978-0-12-386915-9.00016-4.

Lipinski, B., Hanson, C., Lomax, J., Kitinjoa, L., Waite, R., Searchinger, T. (2013) "Reducing good loss and waste," working paper, Installment 2 of Creating a Sustainable Food Future. Washington, DC: World Resources Institute.

Loomes, G. and Sugden, R. (1987) “Testing for regret and disappointment in choice under uncertainty,” Economic Journal, 97(supplement), 118-129.

Lund, C. M., Jaeger, S. R., Amos, R. L., Brookfield, P., and Harker, F. R. (2006) ““Tradeoffs between emotional and sensory perceptions of freshness influence the

price consumers will pay for apples: results from an experimental market," Postharvest Biology and Technology, 41(2), 172-180.

Lundqvist, J., de Fraiture, C. and Molden, D. (2008) "Saving water: from field to fork—Curbing losses and wastage in the food chain," SIWI Policy Brief, SIWI, Stockholm, Sweden.

Milfont, T. L. (2009) "The effects of social desirability on self-reported environmental attitudes and ecological behavior," The Environmentalist, 29(3): 263-269.

Mochon, D. (2013) "Single-option aversion," Journal of Consumer Research, 40(3), 555-566.

Moffatt, P. G. and Peters, S. A. (2004) "Pricing personal services: An empirical study of earnings in the UK prostitution industry," Scottish Journal of Political Economy 51(5): 675-690.

Natural Resources Defense Council [NRDC] (2012) "Wasted: How America is losing up to 40 percent of its food from farm to fork to landfill," Issue paper, August 2012, IP: 12-06-B.

Natural Resources Defense Council [NRDC] (2013) "The dating game: how confusing food date labels lead to food waste in America," NRDC Report, IP: 13-09-A.

Neff, R. A., Spiker, M. L., and Truant, P. L. (2015) "Wasted food: US consumers' reported awareness, attitudes, and behaviors," PloS one, 10(6).

Newsome, R., Balestrini, C. G., Baum, M. D., Corby, J., Fisher, W., Goodburn, K., Labuza, T. P., Prince, G., Thesmarand, H. S. and Yiannas, F. (2014) "Applications and perceptions of date labeling of food," Comprehensive Reviews in Food Science and Food Safety, 13(4): 745-769.

Odermatt, R. and A. Stutzer (2018) "Subjective Well-Being and Public Policy," In E. Diener, S. Oishi, and L. Tay (Eds.), Handbook of Well-Being. Salt Lake City, UT: DEF Publishers.

Østli, J., Esaiassen, M., Garitta, L., Nøstvold, B., and Hough, G. (2013) "How fresh is fresh? Perceptions and experience when buying and consuming fresh cod fillets," Food quality and preference, 27(1): 26-34.

Parfitt, J., Barthel, M., and Macnaughton, S. (2010) "Food waste within food supply chains: quantification and potential for change to 2050," Philosophical Transactions of the Royal Society of London B: Biological Sciences, 365(1554): 3065-3081.

Parizeau, K., von Massow, M., and Martin, R. (2015) "Household-level dynamics of food waste production and related beliefs, attitudes, and behaviors in Guelph, Ontario," Waste Management, 35: 207-217.

Piff, P., Kraus, M., Cote, S., Cheng, B., and Keltner, D. (2010) "Having less, giving more: The influence of social class on prosocial behavior," Journal of Personality and Social Psychology, 99: 771–784.

Piff, P., Stancato, D., Cote, S., Mendoza-Denton, R., and Keltner, D. (2012) "Higher social class predicts increased unethical behavior," Proceedings of the National Academy of Sciences, 109: 4086–4091.

Porpino, G., Parente, J., and Wansink, B. (2015) "Food waste paradox: Antecedents of food disposal in low income households," International journal of consumer studies, 39(6): 619-629.

Porpino, G., Wansink, B. and Parente, J. (2016) "Wasted positive intentions: The role of affection and abundance on household food waste," Journal of Food Product Marketing, 22(7), 733-751.

Principato, L., Secondi, L., and Pratesi, C. A. (2015) "Reducing food waste: An investigation on the behavior of Italian youths," British Food Journal, 117(2): 731-748.

Qualtrics software, Version [December 2016, January 2017] of Qualtrics. Copyright © [2017] Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. http://www.qualtrics.com

Quested, T. E., Marsh, E., Stunell, D., and Parry, A. D. (2013) "Spaghetti soup: The complex world of food waste behaviors," Resources, Conservation and Recycling, 79: 43-51.

Quested, T. E., Parry, A. D., Easteal, S., and Swannell, R. (2011) "Food and drink waste from households in the UK," Nutrition Bulletin, 36(4): 460-467.

Ransom, G. (2005), "Consideration for establishing safety-based consume-by date labels for refrigerated ready-to-eat foods, National Advisory Committee on Microbiological Criteria for Foods," Journal of Food Protection, 68(8): 1761–1775.

Rosen, S. (1986) „The theory of equalizing differences", Ed: Ashenfelter, O. and Layard, R., in: Handbook of Labor Economics, Elsevier Science Publishers BV.

Sagi, A., and Friedland, N. (2007) "The cost of richness: The effect of the size and diversity of decision sets on post-decision regret, " Journal of Personality and Social Psychology, 93(4), 515.

Samotyja, U. (2015) "Influence of shelf life labelling on the sensory acceptability of potato snacks," British Food Journal, 117(1): 222-233.

Sausgruber, R. and Tyran, J-R. (2011) "Are we taxing ourselves? How deliberation and experience shape voting on taxes", The Journal of Public Economics, 95: 164-176.

Schanes, K., Dobernig, K., and Gözet, B. (2018) "Food waste matters - A systematic review of household food waste practices and their policy implications," Journal of Cleaner Production, 182: 978-991.

Schwartz, B., Markus, H. R., Snibbe, A. C. (2006) "Is freedom just another word for many things to buy? ," New York Times Magazine. http://works.swarthmore.edu/fac-psychology/755

Schwartz, S. H. (1977) "Normative influences on altruism," Advances in experimental social psychology, 10: 221-279, Academic Press.

Secondi, L., Principato, L., and Laureti, T. (2015) "Household food waste behavior in EU-27 countries: A multilevel analysis," Food Policy, 56: 25-40.

Sen, S., and Block, L. G. (2008) ""Why my mother never threw anything out": The effect of product freshness on consumption," Journal of Consumer Research, 36(1): 47-55.

Shin, J., and Ariely, D. (2004) "Keeping doors open: The effect of unavailability on incentives to keep options viable," Management Science, 50(5), 575-586.

Sjöberg, L. and Engelberg, E. (2009) "Attitudes to Economic Risk Taking, Sensation Seeking and Values of Business Students Specializing in Finance," Journal of Behavioral Finance, 10(1): 33-43.

Smith, V.L., Williams A.W., Bratton W.K. and Vannoni, M.G. (1982) "Competitive market institutions: Double auctions vs. sealed bid-offer auctions", The American Economic Review, 72(1): 58-77.

Stancu, V., Haugaard, P. and Laehteenmaeki, L. (2016) "Determinants of consumer food waste behavior: Two routes to food waste," Appetite, 96: 7-17.

Stefan, V., van Herpen, E., Tudoran, A. A., and Laehteenmaeki, L. (2013) "Avoiding food waste by Romanian consumers: The importance of planning and shopping routines," Food Quality and Preference, 28: 375-381.

Stenmarck, Å., Jensen, C., Quested, T. and Moates, G. (2016) "Estimates of European food waste levels," Commissioned by the European Commission in the FUSION project.

Swiss Federal Office of the Environment [SFOE] (2012), " Erhebung der Kehrichtzusammensetzung 2012."

Tanner, C., B. Ryf and Hanselmann, M. (2009) „Geschuetzte Werte Skala (GWS), Konstruction und Validierung eines Messinstrumentes", Diagnostica, 55(3): 174-183.

Theotokis, A., Pramatari, K., and Tsiros, M. (2012) "Effects of expiration date-based pricing on brand image perceptions," Journal of Retailing, 88(1): 72-87.

Tonin, M. and Vlassopoulos, M. (2015) „Corporate Philanthropy and Productivity: Evidence from an Online Real Effort Experiment," Management Science, 61(8): 1795–1811.

Trautmann, S. T., and van de Kuilen, G. (2015) "Belief elicitation: A horse race among truth serums," The Economic Journal, 125(589): 2116-2135.

Tsiros, M. and Heilman, C. M. (2005) "The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories," Journal of Marketing, 69(2): 114-129.

Turner, B. M., Rim, H. B., Betz, N. E., and Nygren, T. E. (2012), "The Maximization Inventory," Judgement and Decision making, 7(1): 48-60.

Tversky, A. (1972) “Elimination by aspects: A theory of choice,” Psychological Review, 79, 281–299.

US Department of Justice, Office of Public Affairs (2016) “Goldman Sachs Agrees to Pay More than $5 Billion in Connection with Its Sale of Residential Mortgage Backed Securities,” https://www.justice.gov/, date accessed: 21.01.2017.

Venkat, K. (2011) “The climate change and economic impacts of food waste in the United States,” International Journal on Food System Dynamics, 2(4): 431-446.

Verbeke, W. and Ward, R. W. (2006) “Consumer interest in information cues denoting quality, traceability and origin: An application of ordered Probit models to beef labels,” Food Quality and Preference, 17(6): 453-467.

Vickrey, W. (1961) “Counterspeculation, auctions, and competitive sealed tenders,” The Journal of Finance, 16: 8–37.

Visschers, V. H., Wickli, N., and Siegrist, M. (2016) “Sorting out food waste behavior: A survey on the motivators and barriers of self-reported amounts of food waste in households,” Journal of Environmental Psychology, 45: 66-78.

Wang, L. and Murnighan, J. K. (2014) “Money, emotions, and ethics across individuals and countries,” Journal of Business Ethics, 125(1): 163–176.

Wansink, B. and Wright, A. O. (2006) ““Best if used by…” how freshness dating influences food acceptance,” Journal of Food Science, 71(4): 354-357.

Waste and Resources Action [WRAP] (2015) “Helping consumers reducing food waste, Retail survey 2015,” United Kingdom, ISBN: 978-1-84405-476-3.

Watson, M., and Meah, A. (2013) “Food, waste and safety: Negotiating conflicting social anxieties into the practices of domestic provisioning,” The Sociological Review, 60: 102-120.

Weber, E. U., Blais, E. and Betz, N. E., (2002) "A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors," Journal of Behavioral Decision Making, 15: 263-290.

Williams, H., Wikström, F., Otterbring, T., Löfgren, M., and Gustafsson, A. (2012) “Reasons for household food waste with special attention to packaging,” Journal of Cleaner Production, 24: 141-148.

Wilson, N. L., Rickard, B. J., Saputo, R., and Ho, S. T. (2017) “Food waste: The role of date labels, package size, and product category,” Food Quality and Preference, 55: 35-44.

Wiswall, M. and Zafar, B. (2018) “Preference for the Workplace, Investment in Human Capital, and Gender,” Quarterly Journal of Economics.

World Wildlife Fund [WWF] (2012) “Lebensmittelverluste in der Schweiz – Ausmass und Handlungsoptionen.“

WRAP (2008a) “The Food We Waste,” Banbury, UK.

WRAP (2008b) "Research into consumer behaviour in relation to food dates and portion sizes," United Kingdom.

WRAP (2009a) "Household food and drink waste in the UK," Banbury, UK.

WRAP (2011a) "The water and carbon footprint of household food and drink waste in the UK," Banbury, UK.

WRAP (2011b) "New estimates for household food and drink waste in the UK," Banbury, UK.

WRAP (2017) "Estimates of food surplus and waste arisings in the UK," UK.

Zeelenberg, M. (1999) "Anticipated regret, expected feedback and behavioral decision making," Journal of Behavioral Decision Making, 12, 93-106.

Zeelenberg, M., van Dijk, W. W., Manstead A. S. R., and vanr de Pligt, J. (2000) "On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment," Cognition and Emotion, 14(4), 521-541.

On-line sources:

Clerc, A. and Khamis, H., on behalf of the Fédération Romande des Consommateurs [FRC] (2012) "Enquête gaspillage alimentaire : Date de péremption: douze denrées à l'épreuve du frigo !, " retrieved from : https://www.frc.ch/date-de-peremption-douze-denrees-a-lepreuve-du-frigo/

Exodus (2007) "We don't waste food: A householder survey," retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwixzeCz7tzcAhXFL1AKHTeDAZsQFjAAegQIABAC&url=http%3A%2F%2Fwww.wrap.org.uk%2Fsites%2Ffiles%2Fwrap%2FWe_don_t_waste_food_-_A_household_survey_mar_07.db6802f9.6397.pdf&usg=AOvVaw2YdOAfaTZYpF9GUJ0MlPOT

Federal office for the agriculture [FAO] (2016) "Faits, chiffres, astuces. » Retrieved from: https://www.blw.admin.ch/dam/blw/fr/dokumente/Politik/Food%20Waste/Lebensmittel%20wegwerfen.%20Das%20ist%20dumm.%20Brosch%C3%BCre.pdf.download.pdf/Lebensmittel%20wegwerfen.%20Das%20ist%20dumm._f%20(Brosch%C3%BCre).pdf that takes its sources from the International Food Policy Research Institute.

Rivet-Bonjean, M.-C. (2018) "Quels aliments peut-on consommer sans risque après la date de préemption ? » published on the 8.6.2018 and updated on the 18.6.2018, retrieved on www.lanutrition.fr.

Waste and Resources Action Program (2011) "Consumer insights: Date labels and storage guidance, executive summary," retrieved from: http://www.wrap.org.uk/content/consumer-insight-date-labels-and-storage-guidance.

Curriculum vitae

Fanny Brun
05.05.1988

EDUCATION

08/2014 – 02/2019	**Doctoral Student in Behavioral Economics** Zurich Graduate School of Economics, Zurich, Switzerland
2010 – 2013	**MA in Economics with focus on Econometrics** – Summa cum laude University of Zurich, Switzerland; Korea University, Republic of Korea
2007 – 2010	**BA in Economics** – Rank: 2nd/44 University of Lausanne, Switzerland

PROFESSIONAL EXPERIENCE

01/2018 – 07/2018	**Active member at the Students Sustainability Commission ETH Zurich**
09/2015 – 02/2016	**Research Assistant in Empirical Microeconomics**, University of Zurich Application of quantitative methods to analyze data sets and developing research ideas in the field of Happiness economics.
09/2013 – 08/2014	**Research Assistant in Behavioral Economics and Teaching Assistant in Empirical Methods**, University of Lausanne, Switzerland Designing and conducting an experiment, analyzing the collected data. Responsible for the tutorials in "Empirical Methods in Management and Economics".
05/2013 – 07/2013	**Intern in Quantitative Market Research**, Gesellschaft für Innovative Marktforschung, Heidelberg, Germany Conducting empirical analyses of customers' behavior.
02/2012 – 10/2012	**Event Project Manager and General Secretary**, Golf Club of the University and ETH of Zurich
09/2011 – 07/2012	**Teaching Assistant** at the Chair of Statistics and Empirical Economic Research, University of Zurich
07/2011	**Intern** in the Tax and Legal Services Department, PricewaterhouseCoopers, Geneva, Switzerland
09/2009 – 07/2010	**Research- and Teaching Assistant** at the Chair of Accounting and Control, University of Lausanne, Switzerland
07/2009 – 08/2009	**Intern as Event Project Manager and Researcher** at Alchimie Forever LLC, Washington D.C., USA